Search for neutrinos from annihilating dark matter in galaxies and galaxy clusters with the IceCube detector

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m

(Dr. rer. nat.)

im Fach Physik

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

M.Sc. Meike Machieltje de With

Präsidentin der Humboldt-Universität zu Berlin:

Prof. Dr.-Ing. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:

Prof. Dr. Elmar Kulke

Gutachter: 1. Prof. Dr. Elisa Bernardini
2. Prof. Dr. Marek Kowalski
3. Prof. Dr. Carlos de los Heros

Tag der mündlichen Prüfung: 16.02.2018

Abstract

The identity of dark matter is one of the largest unsolved problems in modern-day (particle) physics. There is a significant number of observations that cannot be explained using current models and dark matter could be a solution for this problem. However, so far all evidence for dark matter is based on gravitational effects. To shed light on the existence and characteristics of dark matter, many experiments are searching for additional evidence. One method for doing this is to search for particles which could have been produced in self-annihilations of dark matter particles.

In this thesis, three years worth of data from the completed IceCube detector is used to search for neutrinos produced in dark matter annihilations in five nearby dwarf galaxies (Segue 1, Ursa Major II, Willman 1, Coma Berenices and Draco), the M31 galaxy and the Virgo cluster. To do this, an event selection which was developed for this analysis is applied to the data sample to reduce the atmospheric background rate from approximately 100 Hz to less than 0.5 mHz. Then, an unbinned maximum likelihood method is used to determine whether there is an excess of neutrinos from the direction of the considered galaxies or galaxy cluster that has an energy spectrum that matches the spectrum expected from dark matter annihilations. For the M31 galaxy and the Virgo cluster an extended signal with a two-dimensional Gaussian shape and width up to 5 degrees is also considered, in addition to the search for a point-like signal.

In all cases, the results of the analysis are compatible with the background-only hypothesis and limits are set on the velocity-averaged dark matter annihilation cross section for different annihilation channels. For the combination of the five dwarf galaxies, these limits (including systematic uncertainties) go down to $3.9 \cdot 10^{-22}$ cm^3 s^{-1} for a dark matter mass of 3 TeV and annihilation to muons. For high dark matter masses there is an excess of neutrinos from three of the five dwarf galaxies. This excess has a global p-value of 4.9%, so it is not statistically significant. The search for an extended emission from the direction of the M31 galaxy and the Virgo cluster also did not result in an excess: in both cases the global p-value is larger than 50%.

The limits on the velocity-averaged dark matter annihilation cross section have improved significantly (up to an order of magnitude) with respect to the previous IceCube analysis considering these same targets. This is partially due to general improvements, for example the increased data sample, larger detector and improved direction reconstruction. In addition, there were improvements to this analysis specifically: an improved event selection was used to select the final data sample and an unbinned maximum likelihood method was used for the final analysis instead of a binned analysis method.

Zusammenfassung

Die Identität der Dunklen Materie ist eines der größten ungelösten Probleme in der modernen Physik. Es gibt eine große Anzahl von Beobachtungen die mann nicht mit aktuellen Modellen erklären kann, und für Dunkle Materie eine Erklärung liefern könnte. Bislang jedoch kommen alle Hinweise für Dunkle Materie aus der Beobachtung von Gravitationseffekte. Um Aufschluss zu geben über die Existenz und die Eigenschaften von Dunkler Materie, suchen viele Experimente nach zusätzlichen Hinweisen. Eine Methode dafür besteht darin nach Teilchen zu suchen, die als Resultat von Selbst-Annihilationen von Teilchen der Dunklen Materie produziert worden sein könnten.

In dieser Dissertation werden Daten aus drei Jahren vom vollständig fertiggestellten IceCube-Detektor benutzt um nach Neutrinos zu suchen , die in Dunkle-Materie-Annihilationen in fünf nahegelegen Zwerggalaxien (Segue 1, Ursa Major II, Willman 1, Coma Berenices und Draco), der M31 Galaxie und dem Virgo Galaxienhaufen produziert wurden. Um die Messung durchzuführen, wurde zunächst eine Ereignis-Selektion angewandt, die es ermöglicht, die von aus der Atmosphäre stammenden Teilchen dominierte Rate der Ereignisse von zirka 100 Hz auf 0.5 mHz zu reduzieren. Danach wird eine Maximum-Likelihood-Schätzer eingesetzt um zu bestimmen ob es ein Überschuss von Neutrinos aus der Richtung der jeweiligen Quellen gibt, der mit einen Energie-Spektrum übereinstimmt das mann für Dunkle-Materie-Annihilationen erwartet. Für die M31 Galaxie und den Virgo Galaxienhaufen wurde zusätzlich zu dieser Suche nach einer Punktquelle auch eine Suche für ein erweitertes Signal mit einer zwei-dimensionalen Gaußschen Verteilung und bis zu fünf Grad Breite durchgeführt.

In allen untersuchten Fällen ist das Ergebnis der Analyse vereinbar mit einer Messung der Hintergrund-Hypothese, und daraus wurden Limits für den über die Geschwindigkeit gemittelten Wirkungsquerschnitt für Dunkle-Materie-Annihilation für verschiedene Endprodukte bestimmt. Für die Kombination von den fünf Zwerggalaxien gehen diese Limits (inklusive systematischer Unsicherheiten) hinab bis $3.9 \cdot 10^{-22}$ cm^3 s^{-1} für Dunkle Materie mit einer Masse von 3 TeV und einer Annihilation in Myonen. Für hohe Dunkle-Materie-Massen gibt es ein Überschuss von Neutrinos aus drei der Zwerggalaxien. Dieser Überschuss hat einen globalen p-Wert von 4.9% und ist damit nicht statistisch signifikant. Die Suche für ein erweitertes Signal von der M31 Galaxie und dem Virgo Galaxienhaufen ergab keinen Überschuss: in beiden Fallen war der globale p-Wert mehr als 50%.

Die Limits auf den über die Geschwindigkeit gemittelten Wirkungsquerschnitt für Dunkle-Materie-Annihilation haben sich im Vergleich zu vorherigen IceCube-Analysen signifikant verbessert, um bis zu einer Größenordnung. Diese ist teilweise auf Grund allgemeine Verbesserungen, wie den großeren Datensatz, großeren Dektor und verbesserte Spurrekonstruktion. Zusätzlich gab es Verbesserungen für diese Analyse besonders: eine verbesserte Ereignis-Selektion um den endgültigen Datensatz zu selektieren, und für die Analyse ist eine Maximum-Likelihood-Schätzer eingesetzt statt eine Analyse in ein Suchfenster.

Contents

1. Introduction

1.1. Dark matter

Gravitational laws allow to determine the force, and thus the acceleration, between two objects once the distance between them and their masses are known. This means that such a law can be used to calculate the behaviour of falling objects on Earth, but also to predict the trajectories of celestial bodies, like the planets in our solar system and stars and gas in other galaxies. Whenever deviations from these predicted trajectories are detected, there are two possible reasons: either the inputs to the gravitational law (the masses and distances) are incorrect, or the gravitational law itself is incorrect.

An example of the first case took place in 1821, when the French astronomer Alexis Bouvard published astronomical tables of Jupiter, Saturn and Uranus based on Newton's law of universal gravitation. The first two matched well with subsequent observations, but the trajectory of Uranus deviated significantly from Bouvard's predictions. He realized that the deviation could be explained by an unknown and so far invisible body. More than 20 years later, John Adams and Urbain le Verrier independently calculated the expected location of this new planet, based on the observed trajectory of Uranus. In 1846 the planet Neptune was found by the German astronomer Johann Galle, very close to the location that Adams and Le Verrier had predicted, which was a triumph for the Newtonian gravitational theory. Though the term was not used at the time, this is a first example of what could be called dark matter: something that was (at first) not detected directly, but was predicted to exist based on gravitational effects.

Only 13 years later, the same Urbain le Verrier reported that the precession of Mercury could not be explained using Newton's law of gravitation and the known planets and Sun. The existence of another planet (named Vulcan) even closer to the Sun was one of the proposed solutions, but this planet was never found. It turned out that Mercury's precession was an example of the second possible cause of deviations: the gravitational law itself was incorrect. The problem was solved by Einstein's theory of general relativity, published in 1915, which simplifies to Newton's gravity law for weak gravitation, static gravitational fields and non-relativistic objects.

In 1933, a Swiss astronomer called Fritz Zwicky also obtained evidence of unseen mass, this time in the Coma galaxy cluster [1]. He estimated the mass of the cluster based on the velocities of its constituent galaxies, and found that this mass estimate was significantly higher than the mass estimate based on the number of galaxies and other luminous matter in the Coma cluster. To solve this problem, he proposed a new component of the galaxy cluster, and referred to this as 'dunkle Materie' (dark matter). Often, this is seen as the birth of the dark matter problem in modern particle physics and cosmology, as Zwicky's 'dark matter' explanation is still the most favoured.

Zwicky's results were the first of many observations of deviations, which will be discussed in the first chapter of this thesis. Just like the earlier deviations in the trajectories of bodies in the solar system, these could either be explained by incorrect input to the gravitational law, or by the gravitational law itself being incorrect. One possible adaptation of the gravitational law, called MOND (Modified Newtonian Dynamics) was proposed by Israeli physicist Mordehai Milgrom in 1983 [2]. He suggested that Einstein's gravitational laws should be adapted in case of accelerations smaller than approximately 10^{-10} m s^{-2}. These accelerations are significantly smaller than accelerations in the solar system and on Earth (the only environments where Einstein's gravitational laws have been tested), but they are typical for galaxy clusters and galaxies.

In general, MOND and similar theories are able to explain some of the current observations, for example the mass discrepancies in galaxies and galaxy clusters. However, so far they have trouble explaining the full spectrum of observations at all scales. For example, observations of the Bullet Cluster can be explained very well by the dark matter hypothesis, while significant fine-tuning is necessary to explain these observations using MOND [3]. For this reason, at the moment dark matter to many people seems to be a better explanation of the observed deviations. However, as long as dark matter has not been observed directly the topic is up for debate, and research into both theories of modified gravity and dark matter will continue. The research described in this thesis is part of the latter group, so for the reminder of this thesis it will be assumed that dark matter exists.

1.2. This thesis

In this thesis, the results of a search for neutrinos originating from dark matter annihilations in nearby galaxies and galaxy clusters using three years of data from the completed IceCube detector are presented. To give a theoretical background, the thesis starts with a thorough description of dark matter in chapter 2, which discusses the evidence for dark matter, the current knowledge about the nature of dark matter (including possible candidates), the different types of dark matter searches and an overview of experimental results. Chapter 3 is devoted to the IceCube detector, including a description of the detection principle, the main experimental backgrounds and the Monte Carlo simulation that is used for the analysis.

The description of the analysis itself starts in chapter 4, which first describes the event reconstruction methods used for this analysis, and then discusses the event selection used to remove as much of the experimental backgrounds as possible. The chapter ends with a description of the data sample that is used for the final analysis. The analysis method itself, an unbinned maximum likelihood analysis, is described in detail in chapter 5, including some of the additional techniques used. Some examples of these are Kernel Density Estimation and an adaptation of the method for extended sources. This chapter ends with the sensitivity of this analysis to the velocity-averaged WIMP annihilation cross section. After applying the event selection and analysis method to the experimental dataset, limits can be calculated, which are the topic of chapter 6. This chapter also discusses the effect of systematic uncertainties and shows comparisons between the limits resulting from this analysis and the limits resulting from other analyses. Finally, chapter 7 contains the conclusion and outlook.

2. Dark matter

Dark matter is one of the most interesting and exciting topics in modern day (astro)particle physics, and its existence is well-motivated, as will be discussed in this chapter. Though some of its properties can be inferred from experimental results, dark matter has never been directly detected and it is not known for certain whether it exists. However, if it exists, this would mean that only about 15% of the total matter density of the University is made up of known types of particles. It would also mean that the current understanding of particle physics (summarized in the Standard Model of Particle Physics) is incomplete, and the properties of dark matter can give important clues to which of the many theories for Beyond the Standard Model physics is the correct one. On the other hand, dark matter is an important part of the current Standard Model of Cosmology. For these reasons, the many aspects of dark matter have been studied in depth, and this chapter will give an overview over the current theoretical and experimental state of the field.

The first section of this chapter describes some of the current evidence for the existence of dark matter, including galaxy rotation curves, gravitational lensing and the Cosmic Microwave Background. In section 2.2, the nature of dark matter is discussed, mentioning the requirements for dark matter candidates and then listing a number of possible candidates, with the main focus on Weakly Interacting Massive Particles (WIMPs). Also, some of the theories which contain WIMP candidates are mentioned briefly. Section 2.3 is devoted to the three types of dark matter searches: direct detection (section 2.3.1), production at collider experiments (section 2.3.2) and indirect detection (section 2.3.3), including recent results. Most of the time, indirect detection experiments are considering a specific target (e.g. the centre of our Galaxy), so after a discussion of dark matter halo models in section 2.4.1, the rest of section 2.4 presents a number of possible targets, and recent experimental results using these targets. Finally, in section 2.5, the properties of the specific signal (neutrinos from dark matter annihilations in dwarf galaxies, the M31 galaxy and the Virgo cluster) searched for in the analysis described in this thesis are listed.

2.1. Observational evidence for dark matter

One of the reasons why the current evidence for the existence of dark matter is so convincing, is that it can be found on different scales. The most well-known and direct evidence comes from studying galaxy rotation curves, which show the rotational velocity of stars and gas in galaxies as a function of the distance to the centre. Following Newtonian dynamics, this velocity is given by

$$v(r) = \sqrt{\frac{GM(r)}{r}} \tag{2.1}$$

with $M(r)$ the total mass enclosed by radius r and G the gravitational constant. This means that at large radii, where (almost) all mass is inside, the rotational velocity should be proportional to $1/\sqrt{r}$. However, from measurements (first done by Rubin and Ford [4], a more recent example is shown in figure 2.1) it is clear that even at radii where there is (almost)

no visible matter, the rotation curve is still flat. This means that either the current theory of gravity is not correct, or there is non-visible (dark) matter extending beyond the visible stellar disk.

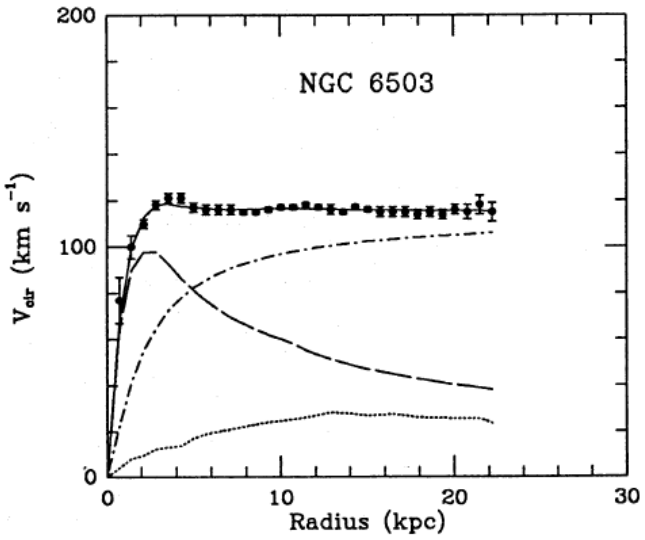


Figure 2.1.: The rotation curve of galaxy NGC 6593, with the dashed, dotted and dashed-dotted curves showing the contributions from the stellar disk, the hot gas and the (presumed) dark halo, respectively. The solid line, which is the sum of these three contributions, fits the data points well. Taken from [5].

On larger scales, evidence for dark matter can be found using gravitational lensing. Following Einstein's theory of general relativity, mass will cause spacetime to be curved [6]. Light following geodesics in this curved spacetime will bend around massive objects, and from the amount of curvature the total amount of mass can be inferred. A good example of this is the 1E0657-558 cluster (often referred to as the 'Bullet cluster') which consists of two clusters that are in the process of colliding [7]. The mass distribution in this area was mapped using the effect of gravitational lensing on the light from sources behind the Bullet cluster [8]. The visible matter in these clusters mainly consists of hot gas, which emits X-ray radiation that was detected by the Chandra satellite [9]. The result can be seen in figure 2.2, from which it is clear that the total matter distribution is very different from the visible matter distribution, which implies that there is a large fraction of dark matter. It also shows that dark matter particles only interact weakly: the hot gas in both clusters is decelerated with respect to the rest of the cluster because of interactions, which is not the case for the dark matter.

On cosmological scales, analyzing the Cosmological Microwave Background (CMB) gives an estimate for the total amount of dark matter in the Universe. The CMB consists of photons that were decoupled from matter shortly after the recombination of electrons and positive ions approximately 377,000 years after the Big Bang, and shows what the Universe looked like at that time [10]. It is isotropic at the 10^{-5} level and follows a black-body spectrum with a temperature of 2.726 K [11]. However, it is the extremely small anisotropies and their angular scales that give an insight into many cosmological parameters.

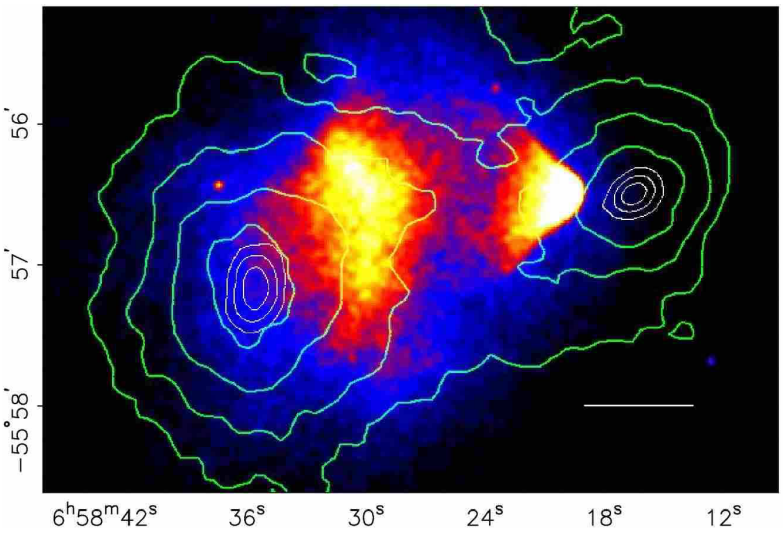


Figure 2.2.: The colliding clusters that form 1E0657-558 (the 'Bullet cluster'), with the X-ray image from the Chandra satellite in colour and the reconstructed weak lensing potential in green contours. The white bar indicates 200 kpc at the distance of the cluster. Taken from [8].

These anisotropies have first been studied by the COBE satellite [12], and the most recent results come from the Planck satellite [13]. These are complemented by a number of ground-based telescopes (mainly the Atacama Cosmology Telescope [14] and the South Pole Telescope [15]) for small angular scales. In figure 2.3 the data points from the Planck satellite are shown in blue, with the best-fit theoretical model shown in red. According to this best-fit theoretical model, the total matter density in the Universe is 31.5% of its total density, while the baryonic matter density is 4.9% of the total density. This means that 26.6% of the Universe must consist of non-baryonic matter, which is expected to be the same dark matter that could explain what is seen at other scales [13]. The other 68.5% of the Universe are believed to be made up of dark energy, about which currently even less is known than about dark matter (see [16] for a review).

2.2. The nature of dark matter

From the previous section, it has become clear that there is significant evidence for the existence of dark matter. Also, a number of its properties are known: it has to be non-baryonic, it has to interact gravitationally, and if it has non-gravitational interactions those have to be weak. This does not mean that dark matter can only interact using the known weak force, it could also interact via a not yet known force of comparable strength. Dark matter also has to be stable or long-lived (with a lifetime much longer than the age of the Universe) to have survived until now.

One candidate that comes to mind immediately are neutrinos, as these are stable fermions which only interact via the weak force, and they are known to exist. However, there are two reasons why neutrinos cannot be all of the dark matter in the Universe. First, their mass density is too small: neutrino masses have not been measured yet, but the current upper limit from a combination of measurements is $\Sigma m_\nu < 0.23$ eV [13]. This means that following the calculation in [17] the neutrino mass density is less than 0.54% of the total density

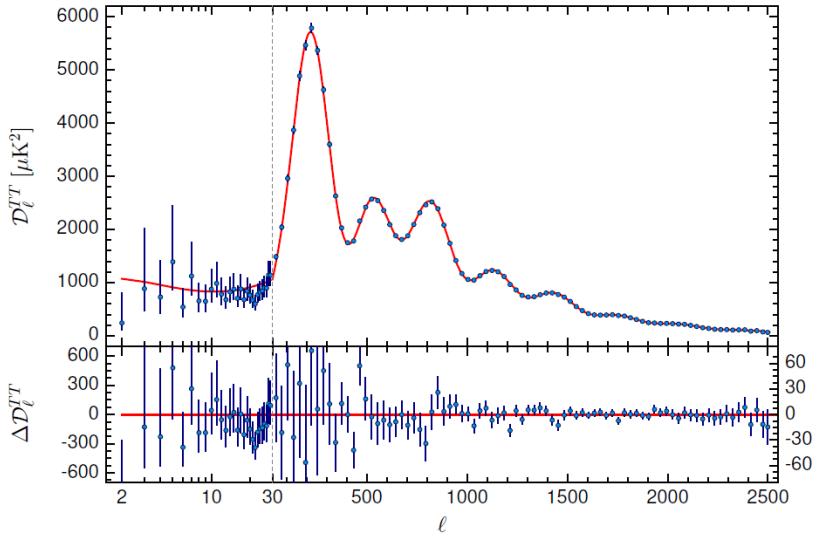


Figure 2.3.: The Planck temperature power spectrum, computed using different methods for multipoles below and above 30 (which lead to the different y scales on the lower plot). The data points are shown in blue, with the error bars showing $\pm 1\sigma$ uncertainties. The best-fit theoretical model is shown in red. Taken from [13].

of the Universe. Second, as neutrinos were relativistic in the early Universe, they would erase density fluctuations at small scales (below approximately $40 \times (m_\nu/(30\text{ eV}))$ Mpc [17]). This means that in the evolution of the Universe, large structures would have formed first and smaller structures only later. Observations strongly suggest this is not the case [18, 19], which means that dark matter is expected to consist of particles that were non-relativistic in the early Universe. Usually, these particles are referred to as cold dark matter.

Another dark matter candidate, which is still an area of active research, is the axion. This hypothetical particle was introduced to solve the problem of strong CP violation in particle physics. This CP violation is in principle allowed, but from measurements it is clear that it must be highly suppressed [20]. A possible solution would be to add a new global symmetry (called a Peccei-Quinn symmetry) that is spontaneously broken, which would result in a new particle: the axion [21]. As axions are very light and very weakly coupled, they are hard to detect. However, they are expected to convert to photons in a strong magnetic field [22]. This is the technique currently used by a number of experiments [23–25], which so far have not detected axions. Even though axions have a very small mass, they are still cold dark matter, as they are already non-relativistic when they are produced [3].

The most widely studied dark matter candidates are Weakly Interacting Massive Particles (WIMPs), which are often denoted by the Greek letter χ. As this thesis describes a search for WIMPs, from now on the terms WIMP and dark matter will be used interchangeably. One of the reasons why WIMPs are so popular is that they are a natural component of a number of theories devised to solve well-known problems with the Standard Model of Particle Physics. Two examples are discussed in sections 2.2.1 and 2.2.2.

In most models, WIMPs are thermal relics, which means that they were created thermally in the early Universe. At that time, when the temperature was high ($T \gg m_\chi$), the

creation and annihilation processes were in equilibrium. As the Universe expanded, it also cooled down, and when $T \approx m_\chi$ the number of WIMPs started to decrease exponentially as $e^{-m_\chi/T}$ [3]. This is because only collisions of particles and antiparticles with relatively high kinetic energy (in the high-energy tail of the Boltzmann distribution) could produce WIMPs. If this had continued, there would be no dark matter left at the current time.

However, due to the expansion of the Universe, at some point the mean free path between WIMP collisions became larger than the Hubble radius, which means that annihilation effectively stopped (except in areas with increased dark matter density). This process is called 'freeze-out', and in the standard cosmological scenario it took place when $T \approx m_\chi/20$ [3]. From then onwards the number density of WIMPs in a comoving volume was approximately constant, and this density is called the relic density. The value of the relic density depends on the annihilation cross section: a higher annihilation cross section means that it took longer until the mean free path was larger than the Hubble radius. In that case, the exponential decrease continued for a longer time, meaning that the relic density is lower. A consequence of this relation is that for standard assumptions and if there is only one WIMP, it is possible to calculate the annihilation cross section from the known WIMP relic density [3].

From detailed calculations it can be shown that the relic density of a particle with weak-scale interactions can be approximated by [17]

$$\Omega_\chi h^2 \approx \frac{3 \cdot 10^{-27}}{\langle \sigma v \rangle}, \tag{2.2}$$

where Ω_χ is the fraction of the total energy density of the Universe that consists of dark matter, h is the Hubble constant in (km/s)/Mpc divided by 100 and $\langle \sigma v \rangle$ is the velocity-averaged WIMP annihilation cross section in $\mathrm{cm}^3\ \mathrm{s}^{-1}$. From this one can calculate that the correct relic density is reached for a velocity-averaged WIMP annihilation cross section of approximately $3 \cdot 10^{-26}\ \mathrm{cm}^3\ \mathrm{s}^{-1}$, which is of the order of the weak interaction.

A more detailed calculation, which for example takes into account the correct number of relativistic degrees of freedom, shows that there is a significant mass-dependence for WIMP masses below 30 GeV [26]. The annihilation cross section is $5.2 \cdot 10^{-26}\ \mathrm{cm}^3\ \mathrm{s}^{-1}$ for a WIMP mass of 0.3 GeV and decreases to $2.2 \cdot 10^{-26}\ \mathrm{cm}^3\ \mathrm{s}^{-1}$ for a WIMP mass of 30 GeV. For masses larger than 30 GeV, the value rises slowly with mass, reaching $2.4 \cdot 10^{-26}\ \mathrm{cm}^3\ \mathrm{s}^{-1}$ for a WIMP mass of 10 TeV.

Usually, WIMPs are assumed to be lighter than approximately 100 TeV. This follows from the fact that using unitarity arguments, one can show that there is an upper limit on the annihilation cross section of thermal relics [27]. This upper limit is mass-dependent and at around 30 TeV [17] crosses the lower limit given by the known relic density of dark matter. However, larger masses are possible for particles that were not in thermal equilibrium during freeze-out, as in that case the dark matter density does not depend on the annihilation cross section, but mainly on the production cross section. These superheavy dark matter particles could for example be produced directly after inflation, in the transition from an inflationary universe to a normal, matter-dominated universe [28]. Superheavy particles would be hard to detect via traditional dark matter detection methods (discussed in section 2.3), but if they were not stable, their decay products would be highly energetic and could be detected on Earth.

2.2.1. Supersymmetry

The most well-known theory for Beyond the Standard Model physics is supersymmetry. It adds a symmetry between bosons (which are the force carriers in the Standard Model) and fermions (which are the constituents of matter) and thus effectively doubles the number of particles [29]. Also, it gives a reason for the extreme difference between the Planck scale and the Higgs boson scale, which is often called the hierarchy problem.

There are many variations of supersymmetry, ranging from the minimal version, where each Standard Model particle has one supersymmetric partner, to more extended versions with many new particles. In most of them, a new conserved multiplicative quantum number called R-parity is introduced, which gives a natural explanation for the long proton lifetime [29]. This is necessary as up to now, proton decay has not been detected (the lower limit on the lifetime is of the order of 10^{32} to 10^{34} years, depending on the decay channel [30–32]). All Standard Model particles have R-parity 1 and all new particles have R-parity -1. This means that when a supersymmetric particle decays, at least one of the decay products has to be another supersymmetric particle, which means that the lightest supersymmetric particle (LSP) is a stable particle, making it an appealing dark matter candidate.

In the minimal version of supersymmetry, there are three possible LSPs which could be dark matter candidates: the sneutrino (the superpartner of the neutrino), the neutralino (a linear combination of the superpartners of the $SU(2)$ weak gauge bosons, the hypercharge gauge boson and the Higgs boson) and the gravitino (the superpartner of the graviton) [3]. The sneutrino has been excluded already by experiments [33]. The gravitino is almost impossible to detect because it only interacts gravitationally, but that means the next-to-lightest supersymmetric particle (NLSP) is meta-stable, which could give some interesting signatures. The neutralino has long been one of the most popular dark matter candidates and is still an active topic of research. However, as no signs of supersymmetry have been found so far, the neutralino is becoming somewhat less popular.

2.2.2. Extra dimensions

The only visible dimensions are three space dimensions and one time dimension, but at higher energy scales other dimensions could appear. This idea is an important part of string theory, but is also explored in its own right, for example in Kaluza-Klein theory [34, 35]. It can solve the hierarchy problem discussed above by assuming that gravity can also propagate in these extra dimensions, which dilutes the strength of gravity that can be observed in the four-dimensional world [3].

Assuming that (some of) the Standard Model fields are free to propagate in the extra dimension, a ‘tower’ of particles with increasing masses is predicted to exist for each of these fields [3]. These new particles are excitations of the Standard Model fields, meaning that they have the same charge and other Standard Model quantum numbers, but a larger mass. If one assumes momentum is conserved in the extra dimension, this leads to a new conserved quantum number called KK-parity. This has the same effect as R-parity for supersymmetry: the lightest KK excitation is a stable particle and as such, a dark matter candidate. This lightest KK excitation can for example be the first excitation of the hypercharge gauge boson, the Z boson or the Higgs boson, and all of these have been studied both theoretically and

experimentally. So far no signal has been detected, and a number of simple models have been excluded already [3].

2.3. Current limits on dark matter

To study the properties of the hypothetical dark matter particle, it is necessary to detect it. There are three main methods to detect dark matter: producing dark matter particles at particle colliders, detecting WIMP-nucleon interactions (direct detection) and detecting the result of dark matter annihilations or decays in the Universe (indirect detection). Each of these methods has different uncertainties and backgrounds, which results in them being complimentary. For an eventual discovery of dark matter particles to be certain and generally accepted, it will be necessary for them to be detected by several experiments, using different methods.

So far, no confirmed dark matter detection has taken place that is accepted within the community, though there are several hints in recent experiments which could in the future turn out to be true detections. In the absence of a detected signal, limits on dark matter models can be placed. This can be done by either studying specific models (for example some of the models discussed above) and excluding them or putting limits on their parameters, or by limiting a more general parameter space, for example the WIMP-nucleon interaction cross section versus the dark matter mass.

2.3.1. Direct detection

The WIMP flux at Earth is of the order of $10^5 \times (100\ \mathrm{GeV}/m_\chi)\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$ [3] under standard assumptions. Even with the small interaction cross section assumed for WIMPs, these WIMPs will sometimes scatter off a nucleus while passing through a detector and cause it to recoil. If the background rate is sufficiently low, this small energy transfer could be detected [36]. A low background rate is usually reached by building the detectors far underground to reduce the number of cosmic rays, implementing a veto around the detectors, and using materials which have low natural radioactivity.

There are different ways to detect the nuclear recoil. One method is to detect phonon excitations of the detector material, which can for example be done by measuring its temperature very precisely. A second method is to detect the electrons or ions produced when a nuclear recoil ionizes atoms, by applying an electric field to the detector material. The third method is to use a scintillating detector material and detect the light that is produced when atoms are excited by the nuclear recoil and then return to their ground state. Usually a combination of at least two of these methods is used, both to filter out noise and because the relative size of the signals is different for nuclear recoils (signal events) and gamma-ray and electron interactions (some of the main backgrounds) [37].

In absence of a signal, direct detection experiments can put a limit on the WIMP-nucleon cross section. The total cross section consists of a spin-independent and a spin-dependent contribution, which can be constrained separately. For spin-independent interactions, the cross section rises with the mass of the target nucleus [36], so detectors using high-mass atoms like xenon give the best result. On the other hand, for spin-dependent interactions the cross section is proportional to $J(J+1)$, where J is the magnitude of the nuclear spin, and

the target nucleus mass does not play a role [36]. Some of the best limits on spin-dependent WIMP-proton interactions come from using the Sun, which consists of protons, as detector material. This is further discussed in section 2.4.5. For both types of interactions, the limits also depend on the WIMP mass: a higher mass means larger energy transfers which are easier to detect, but also a lower WIMP flux.

There are many direct detection experiments currently active, and limits have been set on both the spin-dependent and spin-independent cross section as a function of the WIMP mass [37–39]. The current best limits on both cross sections are shown in figures 2.4 and 2.5. There has also been some excitement regarding a possible dark matter detection by the DAMA experiment. DAMA is different from most other direct detection experiments because it uses only one of the three methods discussed previously: scintillating detector material. Because of this, it is more difficult to remove background and noise events from the event sample, compared to experiments which use multiple methods simultaneously. However, they have observed an annual modulation in their event rate with a significance of 9.3σ [40]. Due to the Earth's rotation around the Sun, moving sometimes against and sometimes with the WIMP flux, one would expect such an annual modulation if some of the detected events were due to dark matter interactions.

The part of the parameter space which would best explain the DAMA observations is shown in figure 2.4, together with possible but low-significance signals in the CoGeNT [41], CRESST-II [42] and CDMS-Si [43] experiments. However, from figure 2.4 one can also see that this part of the parameter space has been excluded by multiple other experiments, including the second phase of the CRESST-II experiment [44] and different analyses of the CDMS data [39]. There are several possible explanations for this: there has been quite some discussion about backgrounds which could have a similar annual modulation to what is seen in the DAMA experiments [45, 46], but there are also non-standard dark matter models in which these results can be reconciled [47].

In general, the limits from direct detection experiments on the WIMP-nucleon interaction cross section depend on several assumptions, which are mostly related to the dark matter distribution in the Solar system. For example, the event rate is directly proportional to the local dark matter density and also depends heavily on the WIMP velocity distribution [3]. For the limits from direct detection experiments in figures 2.4 and 2.5, a local WIMP density of 0.3 GeV cm^{-3} and an isotropic Gaussian velocity distribution are assumed.

2.3.2. Production at colliders

As WIMPs are weakly interacting and either stable or long-lived, in a particle collider they will behave similar to neutrinos: they leave the detector without a trace, since they do not interact with the detector material. These particles can only be detected by considering conservation of momentum: as the total transverse momentum (momentum perpendicular to the beam axis) before the collision is approximately zero, the same must be true for the total transverse momentum after the collision. If this is not the case when the transverse momenta of all visible particles are summed, it is likely that at least one 'invisible' particle was produced in the collision, and it is possible to reconstruct the transverse momentum of this particle. However, it is also possible that one or more particles were produced outside

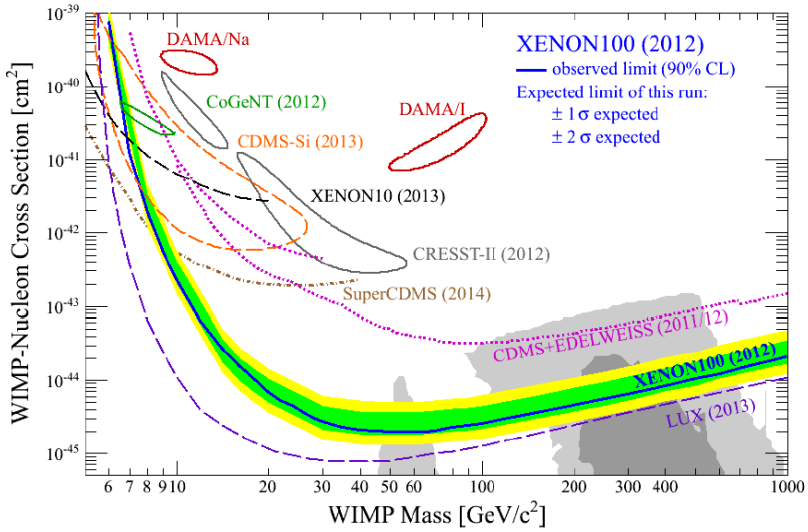


Figure 2.4.: The 90% C.L. exclusion limits for the spin-independent WIMP-nucleon interaction cross section for a number of experiments: XENON10 [48], XENON100 [37], LUX [38], SuperCDMS [39] and CDMS+EDELWEISS [49]. The 2σ detection claims from DAMA [40], CoGeNT [41], CRESST-II [42] and CDMS-Si [43] are also shown. This figure was taken from [50].

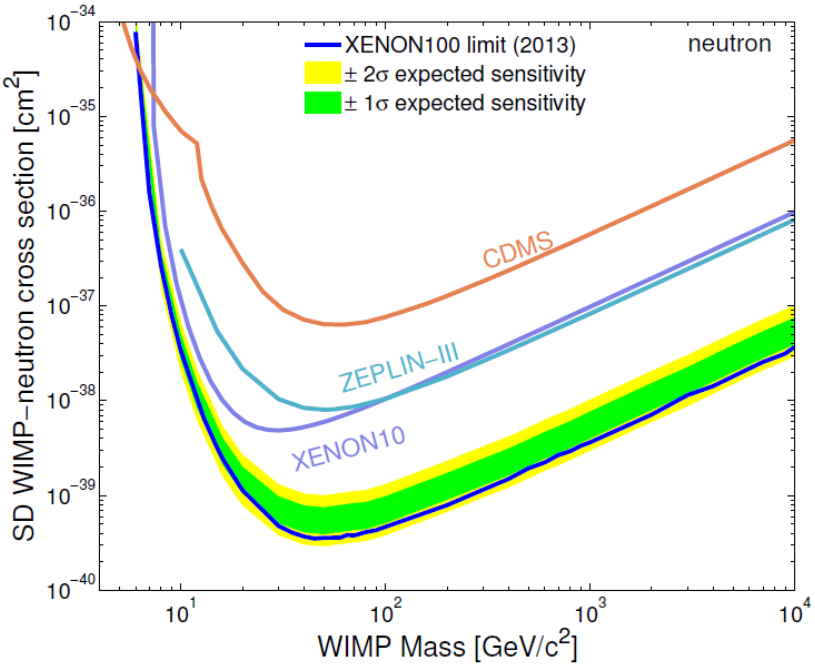


Figure 2.5.: The 90% C.L. exclusion limits for the spin-dependent WIMP-nucleon interaction cross section for a number of experiments: XENON10 [48], XENON100 [37], CDMS [51] and ZEPLIN-III [52]. This figure was taken from [50].

the sensitive area of the detector.

Currently, the particle collider with the highest centre-of-mass energy and luminosity (a measure for the number of collisions per unit time and area) is the Large Hadron Collider

(LHC) in Switzerland. There are a number of detectors which detect the particles that are produced in these collisions, but for dark matter searches the most relevant are the two large all-purpose detectors: ATLAS and CMS.

Traditionally, collider experiments have spent a lot of resources on excluding (parts of the parameter space for) specific theories which contain a dark matter candidate. Using LHC data from Run I (which took place from 2010 until 2013), most of the supersymmetric parameter space where the new particles have masses lower than about 1 TeV has been excluded [53]. Similarly, large advances have been made in excluding theories with extra dimensions [53]. All of these theories could still be true at higher energy scales, but in that case they are no longer a natural solution to the hierarchy problem [29]. LHC Run II has started in 2015 and has a higher centre-of-mass energy than Run I (13 TeV instead of 7 or 8 TeV) so new parts of the parameter space for many models will be explored. It will be interesting to see if anything is detected.

There are also more model-independent ways of searching for dark matter with colliders. In most models, the dark matter particle is the lightest particle for which a new quantum number (for example R-parity in supersymmetry) is non-zero. A consequence of this is that dark matter particles can only be produced in pairs at colliders. In the case of direct production of dark matter particles, there would be no visible particles after the collision (as the dark matter particles themselves are invisible, do not decay and are not expected to radiate off visible particles), so the event would not be considered for further analysis. However, if a visible particle radiated off the beam particle before it collided, this could trigger the detector and be used to calculate the transverse momentum taken away by the invisible particles. Both ATLAS and CMS, the two main experiments at the LHC, have placed limits on the production cross section for dark matter particles using different visible particles, for example gauge bosons [54, 55], jets [56–58], leptons [59] and photons [60, 61].

If one assumes that the particles which mediate the interactions between the WIMPs and the Standard Model (SM) are so heavy that they are not directly accessible to the experiments, all interactions between WIMPs and SM particles can be described in an Effective Field Theory (EFT) framework [62]. In this framework, the mediators are integrated out and interactions between the WIMP and SM particles can be described as contact interactions. These contact interactions are described by a limited number of operators. For each operator, assuming that that operator is dominant, all interactions of WIMPs and Standard Model particles (including production, WIMP-nucleon interaction and annihilation) can be described using only two parameters: the mass of the WIMP and the suppression scale of the heavy mediator particles that are integrated out [62]. This means that if the EFT approach holds, it is possible to directly compare limits from collider experiments, direct detection experiments and indirect detection experiments for a specific operator. This is done for collider experiments and direct detection experiments in figure 2.6.

For direct detection (where the momentum transfer is of the order of tens of keVs) and indirect detection (where the annihilating WIMPs are non-relativistic) it is reasonable to assume that the EFT formalism holds. For collider experiments, on the other hand, it would mean that all new particles in the theory except for the WIMP have masses larger than approximately a few TeV [63]. This is not the case for most realistic models, and in that case the EFT description breaks down. Also, in most models there will be several operators involved

in the interactions between WIMPs and SM particles, and the most relevant operator can be different for the different types of interactions [64]. For these reasons, the usefulness of comparisons between different types of experiments using the EFT formalism is limited, but it is shown here because it is an interesting step towards being able to combine results from different experiments.

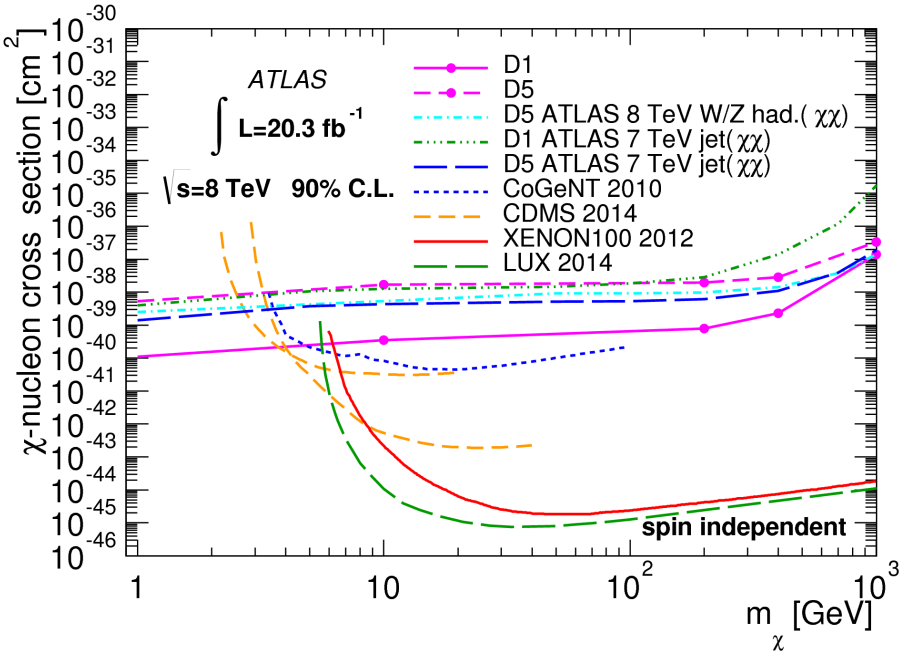


Figure 2.6.: Observed 90% C.L. upper limits on the WIMP-nucleon scattering cross section as a function of the WIMP mass for spin-independent interactions. The pink ATLAS limits are from searches for a Z boson and missing energy [54], and D1 and D5 denote different effective operators. All ATLAS limits shown here assume that the effective operator approach holds, while the limits from the direct detection experiments CoGeNT, CDMS, XENON100 and LUX are independent of this assumption. Taken from [54].

Collider experiments are able to test a part of the parameter space for WIMPs that is hard to access for direct detection experiments, namely WIMPs with low masses (below a few GeV). These WIMPs show only faint signals in direct detection experiments, while in collider experiments the sensitivity depends mainly on whether there is enough energy available to produce a WIMP, so they are sensitive especially at low masses. Astrophysical uncertainties (for example, the uncertainty on the relic WIMP velocity distribution) also do not affect collider experiments. However, colliders cannot test whether a WIMP is stable on time scales much larger than it takes to cross the detector. That means that if a particle is detected, a collider experiment cannot confirm that it is indeed a dark matter candidate (which has to be stable or have a lifetime much larger than the age of the Universe). For that, the other types of experiments are necessary.

2.3.3. Indirect detection

There are two ways in which dark matter can produce stable particles which can be detected on Earth: via self-annihilations and via decays. In most models, the WIMP is a stable

particle (usually because a new conserved quantum number has been introduced) and only annihilations are possible. However, in some models, for example supersymmetric models with R-parity violation (see [65] for an overview) the conservation of this new quantum number is violated by a small amount. This allows WIMP decays to take place at low rates.

In both self-annihilations and decays, the products will in most cases be Standard Model particles, for example gauge bosons or quarks. In the decay process which follows, stable Standard Model particles (photons, neutrinos, electrons, protons and their antiparticles) will be produced, which could travel through the Universe and be detected on Earth.

For annihilations, the expected flux of Standard Model particle type i is given by [3]

$$\frac{d\phi_i}{dE} = \frac{\langle \sigma_A v \rangle}{4\pi \cdot 2m_\chi^2} \frac{dN_i}{dE} J(\Delta\Omega) \tag{2.3}$$

with $\langle \sigma_A v \rangle$ the velocity-averaged annihilation cross section, m_χ the WIMP mass, dN_i/dE the expected energy spectrum per dark matter annihilation and $J(\Delta\Omega)$ the so-called (integrated) J-factor. This is defined as

$$J(\Delta\Omega) = \int_{\Delta\Omega} d\Omega \int_{l.o.s.} \rho^2(l) dl \tag{2.4}$$

where l is the coordinate along the line of sight, ρ is the dark matter density and $\Delta\Omega$ is the solid angle being integrated over. From this equation it is clear that the J-factor not only depends on the total amount of dark matter in the line of sight and the distance to the Earth, but also on how the dark matter is distributed. This will be discussed in detail in section 2.4.1.

For decays, the formula for the flux is rather similar to that for annihilations [3]:

$$\frac{d\phi_i}{dE} = \frac{1}{\tau} \frac{1}{4\pi m_\chi} \frac{dN_i}{dE} J_d(\Delta\Omega) \tag{2.5}$$

with τ the WIMP lifetime and $J_d(\Delta\Omega)$ the (integrated) J-factor for decay, which is defined as

$$J_d(\Delta\Omega) = \int_{\Delta\Omega} d\Omega \int_{l.o.s.} \rho(l) dl. \tag{2.6}$$

In equation 2.5, the $1/\tau$ takes the place of the $\langle \sigma_A v \rangle$ in equation 2.3. The lack of a factor 1/2 and the replacement of ρ^2 by ρ is due to the fact that only one WIMP is necessary for decay, instead of two for self-annihilation. As the topic of this thesis is a search for annihilating dark matter, from now on only this type of dark matter will be considered. However, many of the methods discussed here can also be applied to decaying dark matter.

The expected energy spectrum dN_i/dE depends on the branching ratios to the different annihilation channels, which are model-dependent. To make the limits more model-independent and make it easier to compare between different experiments, usually a number of benchmark channels are chosen. For each of these a limit on the annihilation cross section is calculated assuming a 100% branching ratio to that channel. The limit on the annihilation cross section for a specific model with known branching ratios will then be somewhere in between these benchmark limits. Like for direct detection, the limit will depend on the WIMP mass, as for lower masses the annihilation products will have lower energies and thus be harder to

detect, while for higher masses the annihilation rate goes down because there are less WIMPs.

In general, the current velocity-averaged annihilation cross section $\langle \sigma_A v \rangle$ is equal to the velocity-averaged annihilation cross section in the early Universe [3]. This would mean that for a typical thermal relic WIMP one would expect $\langle \sigma_A v \rangle$ to be approximately $3 \cdot 10^{-26}$ cm^3 s^{-1}, as seen in section 2.2. However, experiments which are only sensitive to higher values of $\langle \sigma_A v \rangle$ are still interesting: as discussed in section 2.2, there are models where the WIMP is not a thermal relic and thus there is no reason to assume that $\langle \sigma_A v \rangle$ should be $3 \cdot 10^{-26}$ cm^3 s^{-1}. It is also possible that the current $\langle \sigma_A v \rangle$ is not equal to that in the early Universe. In most models, the annihilation cross section does not depend on the relative velocity of the two WIMPs, but under some circumstances this is not the case. Since the relative velocity of two annihilating WIMPs in for example a current-day galaxy is $\sim 10^{-3}$, while it was about 0.27 at the time of freeze-out [3], this can have a large effect.

One example of a case where the annihilation cross section depends on the relative WIMP velocity is the Sommerfeld effect. This can take place if dark matter interactions are mediated by a force carrier ϕ with a mass much lower than that of the dark matter particle. In that case, the force carrier can be exchanged between the two WIMPs multiple times in the annihilation process, which can enhance the annihilation cross section, depending on the relative velocities of the WIMPs [66]. The enhancement is larger if the interaction between ϕ and the WIMP is stronger, however in that case the annihilation rate of WIMPs into ϕ particles will also be higher, which eventually leaves one with not enough dark matter to explain the missing matter in the Universe. Detailed calculations in [67] show that in minimal models, the maximum enhancement is of the order of a factor 100.

In the remainder of this section, the three main channels for indirect detection will be discussed: charged cosmic rays, photons and neutrinos.

Charged cosmic rays

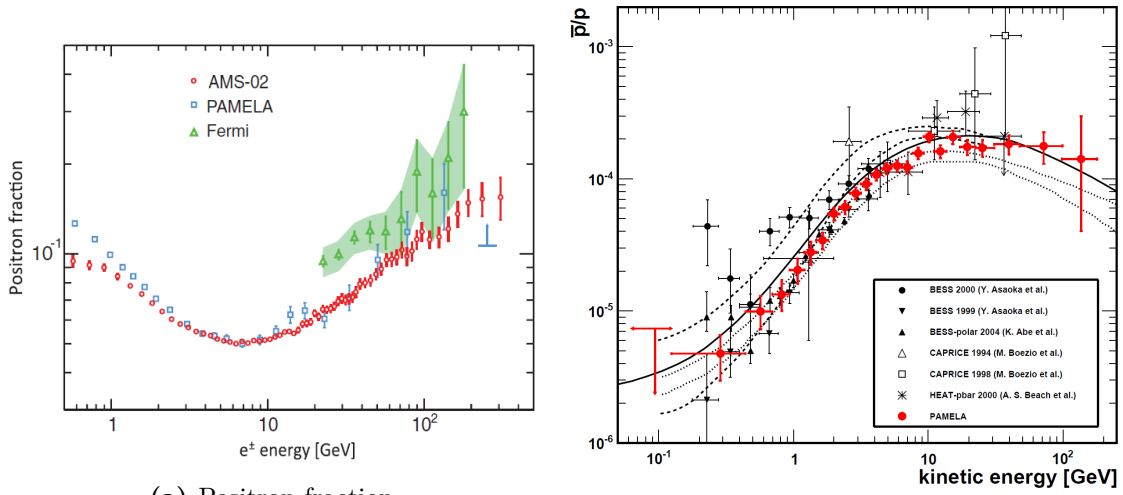
During their propagation from the production site to the detector, charged cosmic rays will be deflected by Galactic magnetic fields. Due to the lack of knowledge about the exact orientation and strength of these fields, it is currently not possible to reconstruct the path of charged cosmic rays [3]. Therefore, only the total flux can be detected, not the flux from any specific target, and the direction of a charged cosmic ray cannot be used to separate signal and background. The expected total flux on Earth also depends on the specifics of propagation of charged particles in the Galaxy, which are not known to great detail [68].

Due to conservation of quantum numbers like charge and lepton numbers, equal numbers of particles and antiparticles will be produced in dark matter annihilations. When looking for dark matter annihilation using charged cosmic rays, usually only antiparticles (positrons, anti-protons and anti-deuterons) are considered, since the background in these channels is much smaller. For a long time it was assumed that the only non-dark-matter process which produces antiparticles was spallation of nuclei in the interstellar medium by the impact of cosmic rays [3]. However, it turns out that this is not true, as will be discussed later on.

Since charged cosmic rays tend to interact in the Earth atmosphere, the current generation of detectors mainly consists of satellite experiments. These can detect the velocities and energies of incoming cosmic rays and identify them before they hit the atmosphere. In most

cases, studying the antiparticle fraction instead of the total antiparticle flux is less complicated, since many uncertainties cancel out in this case.

If spallation is the only source of positrons, the positron fraction is expected to fall with increasing energy. The current best results are shown in figure 2.7a and come from the Alpha Magnetic Spectrometer (AMS-02) [69] located on the International Space Station. This experiment has confirmed [70] earlier results [71–73], where after an initial fall, a rise in the positron fraction starting at about 7 GeV was seen. This means that there must be another source of positrons nearby (since charged particles lose energy in magnetic fields, they cannot have travelled far).



(a) Positron fraction

(b) Antiproton fraction

Figure 2.7.: Measurements of the positron fraction (a) and antiproton fraction (b). (a) shows the positron fraction as a function of the positron/electron energy, these are results from the AMS-02 [70], PAMELA [72] and Fermi-LAT [73] experiments. This figure was taken from [70]. (b) shows the antiproton fraction as measured by PAMELA and several older experiments, compared to theoretical calculations for a pure secondary production of antiprotons during the propagation of cosmic rays in the galaxy. This figure was taken from [74].

Annihilating or decaying dark matter would be a possible explanation for the positron excess, but not a natural one. The necessary annihilation cross section is several orders of magnitude above the thermal relic cross section, and to not be in conflict with experimental limits from antiproton searches the dark matter would have to be leptophilic [75].

Another possible source of high-energy positrons would be nearby pulsars [68], which are rotating neutron stars with strong magnetic fields. In these fields, electrons can be accelerated to high energies and induce electromagnetic cascades, and the high-energy photons produced there can create high-energy positrons and electrons via pair production. It is complicated to distinguish between the pulsar and the dark matter scenario by only considering the positron fraction as a function of energy. As all detailed studies eventually run into the problem that many of the aspects of positron propagation in the Galaxy are currently not well-known [68], more information about this could possibly break the degeneracy.

The second antimatter channel that is usually considered is antiprotons. The most current result [74] comes from PAMELA, a satellite detector similar to AMS-02. As can be seen from figure 2.7b, the antiproton ratio matches well with theoretical predictions over the whole energy range. However, like for positrons, the theoretical uncertainties related to propagation of the antiprotons in the Galaxy are relatively large. It would be interesting to see if prediction and experiment still match once the theoretical predictions are more precise.

A third option are searches for anti-deuterons, which are produced in the fusion of antiprotons and anti-neutrons. In contrast to all other channels, here dark matter annihilations are expected to produce an excess at low energies instead of high energies. At the energies that current experiments are sensitive to, the expected background flux is so low (a factor 10^4 lower than the antiproton flux [3]) that less than one background event is expected [76]. For this reason, a single event would be already be interesting, but so far, none have been detected.

Gamma rays

The second interesting channel for indirect detection of dark matter are photons, specifically gamma rays (photons with energies larger than approximately 100 keV). Since gamma rays travel in straight lines, it is relatively easy to determine their source. Another advantage of gamma rays when compared to charged cosmic rays is that for the typical travel distances and energies considered in dark matter analyses, the energy loss during propagation is minimal.

There are two types of gamma-ray detectors currently active: a number of ground-based Imaging Air Cherenkov Telescopes (IACTs) and one satellite detector which is called the Fermi Large Area Telescope (Fermi-LAT) [77]. Fermi-LAT has been in orbit around the Earth since June 2008 and detects gamma rays by measuring the tracks of the electrons and positrons created in pair production in the detector material. It is sensitive to gamma rays from approximately 20 MeV to 300 GeV and to about 20% of the sky at any one time, but due to its continuous motion it covers the whole sky every three hours [77].

At higher energies, the gamma ray flux is so low that to detect a significant number of gamma rays, the effective area of a detector needs to be so large that space-based detectors are no longer feasible. In this regime, the most sensitive detectors are IACTs, which can detect the Cherenkov light radiated by the charged particles produced when a high-energy gamma ray interacts with the Earth atmosphere. Currently, the most sensitive experiments are HESS [78], MAGIC [79] and VERITAS [80]. Typically, these experiments are sensitive to gamma rays from approximately 100 GeV to 100 TeV, and since the atmosphere itself is used as a detector, low fluxes can be detected. However, the typical field of view is only a few degrees, so in general these telescopes are pointed at targets from which significant fluxes are expected and are not used for large sky surveys.

For gamma rays from dark matter annihilations, the energy spectrum depends on the annihilation channel. One interesting annihilation channel is direct annihilation to photons, as the energy of each photon will be equal to the WIMP mass. The resulting line spectrum would be visible as a clear peak in the gamma-ray spectrum to detectors with a good energy resolution. Since it is complicated to explain such a peak with any other background or astrophysical source, this would be a so-called ‘smoking gun’ signal for dark matter annihilation [3]. However, as WIMPs are electrically neutral, they can only annihilate directly to

gamma rays via loop diagrams, which means that this channel is severely suppressed in most models.

All other annihilation channels will produce a continuum spectrum of gamma rays, produced in the decay of primary annihilation products (mainly from neutral pions decaying to photons), internal Bremsstrahlung (photons radiating off virtual particles) and other processes [3]. The spectrum does not contain any clear features except for a cut-off at the WIMP mass, is similar for most annihilation channels, and is also rather similar to that of many types of background.

Due to the good directional resolution, gamma ray searches for annihilating dark matter usually consider specific targets, and the spatial distribution of events is the main property used to distinguish signal and background. In section 2.4, results from gamma ray searches using several types of‘ targets are discussed and compared.

Neutrinos

Like gamma rays, neutrinos are produced in the decay chain of (almost) all primary annihilation products and travel in straight lines, so they point back to their sources. Neutrinos do not lose energy while travelling through the Universe, even over large distances, except when they have extremely high energies. They can also escape dense environments. However, using neutrinos also has a disadvantage: since neutrinos are weakly interacting, large detectors are necessary to be able to detect typical neutrino fluxes. In general, the direction and energy resolution of current neutrino detectors are also worse than those of gamma ray detectors. Since the topic of this thesis is a search for neutrinos produced in dark matter annihilations, chapter 3 is dedicated to discussing neutrino detectors, including the detection principle and main backgrounds.

The expected neutrino energy spectra will be discussed in more detail in section 2.5. The general behaviour is similar to that for gamma rays: direct annihilation to neutrinos will produce a peak at the WIMP mass, and other annihilation channels result in a continuum spectrum. However, it is important to note that the annihilation channels which give the best sensitivity are different. For example, neutrino detectors are more sensitive to annihilation to muons than to annihilation to b quarks, while for gamma ray detectors it is the other way around.

2.4. Promising targets for indirect detection of dark matter

As discussed in the previous section, both neutrinos and gamma rays point back to their sources, so it makes sense to consider specific targets where the expected flux from dark matter annihilations is high. As can be seen from equation 2.3, the flux is proportional to the dark matter density squared, so the most promising targets for dark matter detection are those which are expected to have high dark matter densities. In this section, some specific targets will be discussed. The section will start with an overview of dark matter halo models, as for some targets (like the Galactic Centre) the halo model has a large impact on the expected flux.

2.4.1. Dark matter halo profiles

Rotational velocity curves for galaxies and galaxy clusters (e.g. figure 2.1) can be explained if these objects are embedded in a dark matter halo. Using these curves, one can also approximately calculate the total halo mass and how it depends on r at large distances from the centre. However, to be able to calculate the J-factor (equation 2.4) with any degree of precision, a better knowledge of the dark matter distribution is necessary, especially close to the centre of the distribution. Unfortunately, it is complicated to determine the dark matter density profile in this region using velocity curves, even when studying our own Galaxy [81].

For the past decades, dark-matter-only N-body simulations have been used to study the density profiles of dark matter haloes of different sizes. In these simulations, the evolution of a system of dark matter 'particles' under the influence of gravitational forces is tracked. Because of computational constraints, it is not possible to simulate the actual dark matter particles, so for current galactic halo simulations the 'particle' mass is of the order of 10^3 solar masses. The results of these simulations suggest that haloes at different scales follow similar density profiles, largely independent from the initial conditions and halo mass [82]. These profiles can be well described by a double power law, the so-called Navarro-Frenk-White (NFW) profile [82]

$$\rho_{NFW}(r) = \frac{\rho_0}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2} \tag{2.7}$$

with r the distance from the halo centre, ρ_0 the characteristic density (the normalization factor) and r_s the characteristic radius, which gives the approximate distance where the profile becomes steeper. In figure 2.8, the NFW profile is compared to two profiles that will be discussed later on: the Burkert profile and the Einasto profile. For this comparison, best-fit values for the Galactic halo from [83] (for the NFW and Burkert profiles) and [84] (for the Einasto profile) are used. For the NFW profile, the best-fit value for r_s is 16.1 kpc, and the turnover from a shallower to a steeper profile at this point is indeed visible.

Newer simulations with higher resolutions show that the slope of the central part of the halo might be less pronounced than the NFW profile prediction [85], and that the Einasto profile [86] might be a better fit (see figure 2.8). This is defined as

$$\rho_{Ein}(r) = \rho_s \cdot e^{-\frac{2}{\alpha}((r/r_s)^\alpha - 1)} \tag{2.8}$$

with r_s the radius of the sphere which contains 50% of the total mass, ρ_s the mass density at r_s and α a positive number which describes how steep the profile is. However, it might be that the Einasto profile fits the dark matter distribution better because it has three free parameters (r_s, ρ_s and α) while the NFW profile has only two (ρ_0 and r_s).

In recent years, improvements in detection techniques have made it possible to study the rotation curves of other galaxies in more detail [87–89]. The dark matter profiles that can be inferred from this are often not completely in agreement with those found in N-body simulations: while profiles like the NFW profile and Einasto profile show a steep 'cusp' in the centre of the halo, the density profiles derived from rotation curves are flatter (which is usually called a 'core'). One example of a cored dark matter profile is called the Burkert profile [90] and defined by

$$\rho_{Bur}(r) = \frac{\rho_0}{(1+\frac{r}{r_s})\cdot(1+(\frac{r}{r_s})^2)} \tag{2.9}$$

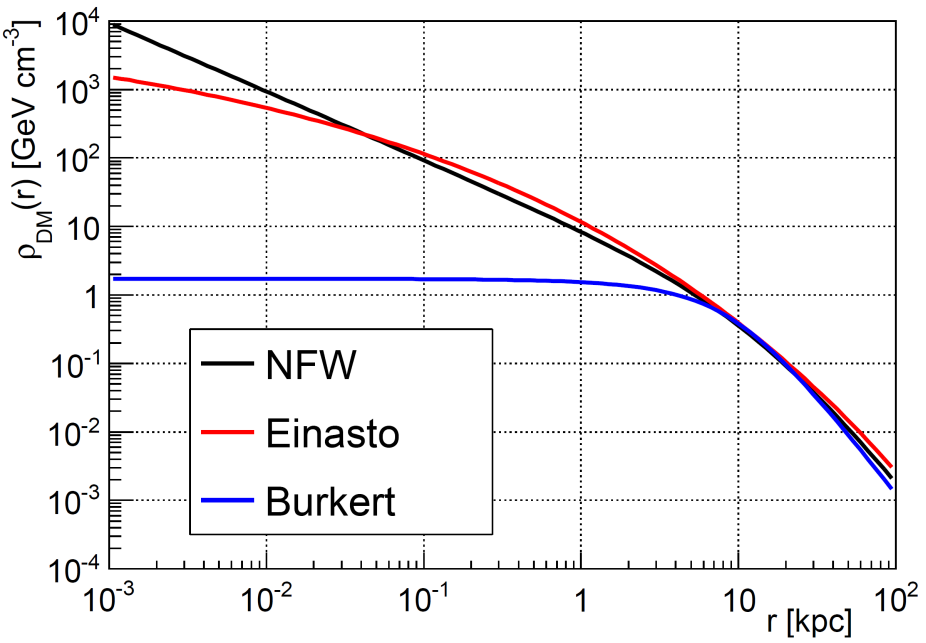


Figure 2.8.: Comparing $\rho_{DM}(r)$ for the NFW, Einasto and Burkert profiles discussed in the text, using best-fit values for this Galaxy taken from [83] (for the NFW and Burkert profiles) and [84] (for the Einasto profile).

with ρ_0 the central density and r_s a scaling radius. From figure 2.8 one can see that the difference between the Burkert profile and the NFW profile is large in the centre of the Galaxy.

The disagreement between the results of N-body simulations and the results of observations is called the cusp-core-problem. It has spawned many discussions about the validity of the cold dark matter model, including alternative theories like warm dark matter [91,92] and self-interacting dark matter [93,94]. However, the simple versions of these alternative theories come with their own consistency problems, both within the model and with observations (for example of the Bullet cluster [95]). On the other hand, more complicated models for dark matter self-interactions [96] so far seem to be consistent with all data.

Another possible explanation for the cusp-core problem is the fact that the N-body simulations are based on incorrect assumptions. In these simulations only dark matter is taken into account. This is done because baryonic physics is much more complicated than the simple gravitational interactions between dark matter particles, and outside the central region, dark matter is the dominant component in galaxies. However, especially in the centre of a galaxy, baryonic physics might have a significant effect on the dark matter distribution. This means that it is necessary to take this into account before making any final statements about cold dark matter being inconsistent with observations.

In the past years, a number of simulations (e.g. [97–99]) have been done where baryonic physics, including the formation of stars and galaxies inside the dark matter haloes, is (partially) taken into account. Some simplifying assumptions, for example that the baryons form an ideal gas and that magnetic fields and radiation can be ignored, are necessary and (at least in the case of the ideal gas assumption) well-motivated [100]. Due to the large number

of adjustable parameters in these simulations and the large effect that certain assumptions can have on the final result, there is no consensus yet on how baryonic physics alters the results of dark-matter-only N-body simulations. However, it is clear that it is possible for cold dark matter combined with baryonic physics to produce both cored and cusped dark matter haloes, and that the assumption that dark matter haloes look the same at all scales is not necessarily true [100].

Substructure

As mentioned before, from observations it is known that the smaller structures in the Universe have formed first, and later on these small structures began to merge. Due to this history, one would expect current dark matter haloes to contain some substructure, and this is also seen clearly in N-body simulations [101]. For Galaxy-sized haloes, approximately 10-50% of the total dark matter is expected to be part of a subhalo, and some of these subhaloes are expected to contain smaller satellite galaxies (dwarf galaxies) [102].

Since the expected flux from dark matter annihilations is proportional to the square of the dark matter density, substructure will boost the expected annihilation signal from a halo. This is usually described in terms of the boost factor, which is the factor with which one needs to multiply the flux from the smooth halo to get the total flux from the halo with substructure. The value of this boost factor is highly dependent on the specific spatial and mass distributions and dark matter profiles of subhaloes. Unfortunately, these are hard to determine, especially for the low-mass end of the spectrum which contributes most to the boost factor. The reason for this is that due to computational limitations it is impossible to simulate each dark matter particle separately, so 'particles' of approximately $10^3 M_\odot$ or larger are simulated instead. With this, the lowest scale at which clustering (the formation of subhaloes) can be detected is currently about $10^5 M_\odot$. For WIMPs, the true limit is expected to be much lower, but the exact value can range from 10^{-4} to $10^{-12} M_\odot$ [103], which can have a large effect on the predicted boost factor [104]. In addition, it is not clear how exactly the subhalo mass distribution should be extrapolated, and as it is being extrapolated over 9 to 17 orders of magnitude this introduces large uncertainties.

Because of all these uncertainties, estimates for the total boost factor vary wildly. Early calculations resulted in boost factors of the order of 1000 for galaxy clusters [104], but more modest estimates (boost factors for galaxy clusters of around 30) [105] are more in agreement with what is seen in recent simulations. For smaller objects, the boost factors are also expected to be significantly smaller, for example for the M31 galaxy the boost factor following [105] would be about 14 and following [104] it would be 66. Even the dwarf galaxies, which themselves are subhaloes of the main halo, are expected to contain substructure (subsubhaloes of the main halo) and could have a boost factor of about 2-3 [105]. However, boost factors are determined from simulations of field haloes (not subhaloes), and tidal stripping could have removed most of the substructure [105]. For this reason, boost factors for dwarf galaxies are almost never considered in experimental analyses, and boost factors for larger objects usually only with many caveats, since their values are so uncertain.

A second effect of substructure is that the dark matter signal would be more extended, which could be visible for dark matter haloes that are both close and large. This is due to the fact that while the dark matter in the halo is highly concentrated towards the centre, that is not the case for the distribution of subhaloes inside the main halo. At relatively large

distances from the centre, there is still a significant number of subhaloes [106]. This will change the spatial profile of the expected signal, but the exact profile strongly depends on model assumptions.

The final effect of substructure is that subhaloes could be detectable by themselves, even if they contain no stars. It is likely that the Galaxy contains subhaloes which are not sufficiently large to capture enough baryons to have star formation, but could still be large enough to have significant dark matter annihilation rates. If one of these subhaloes happens to be close to the Earth, it would be a gamma-ray source with no counterpart at other wavelengths. So far, Fermi-LAT has detected hundreds of these sources [107], though in most cases these are probably known astrophysical sources like blazars and pulsars where the radiation in other wavelengths is not strong enough to be visible. It has been studied whether any of these unidentified sources match the expected energy spectrum for annihilating dark matter, but no evidence has been found so far [108].

2.4.2. The Galaxy

The Galaxy is one of the most interesting targets for indirect detection of dark matter, as the expected dark matter density is high and it is (apart from the Sun and Earth) the closest target. One method is to consider the Galactic Centre, or the area close to the Galactic Centre, where the dark matter density is highest. However, there are many other astrophysical processes taking place in this region, and it is complicated to separate the dark matter signal from these astrophysical processes, especially for gamma ray experiments (where indeed a large background has been detected already [109, 110]). Another tactic is to consider the Galactic Halo far away from the centre, where the signal expectation is lower, but the background is also much lower.

In both cases, an extended signal is sought for, whose exact shape can be calculated for an assumed dark matter profile using equation 2.4. At large distances from the centre of the Galaxy, the shape of the dark matter profile is similar for the different profiles discussed in section 2.4.1, and the normalization is constrained by the fact that the dark matter density in the solar neighbourhood is known reasonably well [3]. As a consequence, the expected signal from parts of the Halo far from the Galactic Centre does not depend greatly on the profile considered. This is not the case if the Galactic Centre region itself is considered. There, the dark matter density can differ by orders of magnitude for different profiles, as can be seen in figure 2.8.

There is no consensus yet on which profile describes the dark matter halo of the Galaxy best. Both cuspy (e.g. the NFW profile) and cored (e.g. the Burkert profile) give reasonable fits to the available data, though at the moment there seems to be a slight preference for cored profiles [83]. In most experimental searches, this problem is solved by quoting limits for both cuspy and cored halo profiles.

Both gamma ray and neutrino experiments have searched for particles produced in dark matter annihilations in the Galactic Centre region. For gamma rays, at high WIMP masses the best limits (for the Einasto profile and the NFW profile) are from the HESS detector [111]. This analysis did not consider the Galactic Centre itself but a region close to it, to avoid most of the background. At lower WIMP masses, Fermi-LAT is expected to give the best limits,

though due to the complexity of the background subtraction the Fermi collaboration itself has not published a limit yet. However, a number of other groups have analyzed the (freely available) data from the Fermi-LAT telescope and found a spherically distributed excess at energies around 1-3 GeV. This is most compatible with a 30-50 GeV WIMP annihilating to b quarks, but other combinations of WIMP masses and annihilation channels also give a good fit [110, 112, 113]. However, there are also other possible explanations, for example a large population of unresolved millisecond pulsars [114, 115] or a recent leptonic cosmic-ray outburst [116]. Currently, the nature of the gamma-ray excess at the Galactic Centre is still an open question. It will likely remain so for some time, though the observation or exclusion of an excess at similar WIMP masses and annihilation cross sections in other targets (most importantly, dwarf galaxies) would certainly be a strong argument for or against dark matter.

As mentioned in section 2.3.3, a gamma ray line would be a clear sign of dark matter, though in most models it is highly suppressed. The Galactic Centre is arguably the best place to look for such a signal: it is the target from which the highest flux is expected (the J-factor is at least one order of magnitude larger than that for other targets), and the high background is less problematic with such a significant difference between the background and signal energy spectra. Recently, there was some excitement about an indication for a 130 GeV gamma ray line in Fermi-LAT data in the direction of the Galactic Centre with a global significance of 3.2 σ [117]. However, after adding more data and improving the event selection, the significance of this potential line was reduced, and no other evidence for a line was found in the energy range from 200 MeV to 500 GeV [118]. The HESS experiment did a similar search at higher energies (from 500 GeV to 25 TeV), but also did not find a significant excess [119].

In addition to the searches for gamma rays, the neutrino experiments IceCube [120] and ANTARES [121] have also analyzed the region of the Galactic Centre and published limits on the annihilation cross section. In figure 2.9, these limits are compared to the HESS limit for an NFW profile and annihilation to b quarks, and it can be seen that for WIMP masses above approximately 500 GeV, the HESS limit is orders of magnitude better than the limits from neutrino experiments. However, one should take into account that for HESS, the limit on annihilation to b quarks is one of the strongest (compared to other annihilation channels), while for the neutrino experiments it is one of the weakest.

IceCube has also published limits on the annihilation cross section based on an analysis of the Galactic Halo using a multipole method to search for a large-scale anisotropy in the neutrino flux [122]. This limit is also shown in figure 2.9.

2.4.3. Dwarf spheroidal galaxies

It has been known for a long time that the Galaxy has a number of satellite galaxies, at distances ranging from about 15 kpc to 250 kpc. Most of these are so-called dwarf spheroidal galaxies (often shortened to 'dwarf galaxies'), which are galaxies much smaller than the Galaxy (typical masses are about 1% of the mass of the Galaxy) with low luminosities. They contain no gas (as far as is currently known) and only a small number of stars, but they are expected to contain a large amount of dark matter, with mass-to-light ratios hundreds of times larger than that for the Sun [123]. For these reasons, they are interesting for dark matter searches: the expected signal flux is relatively high (though the flux expected from

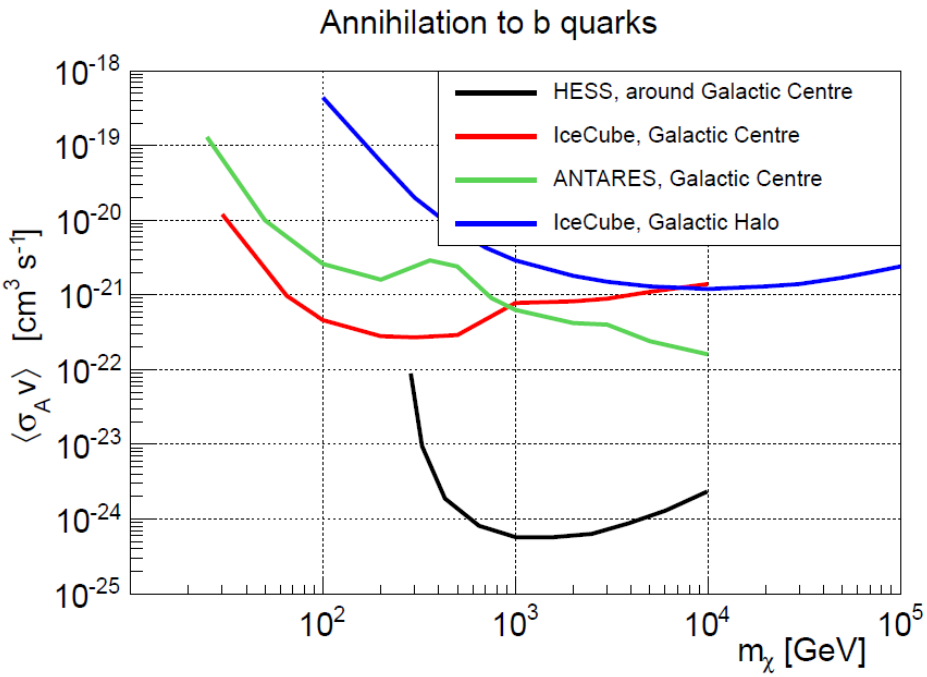


Figure 2.9.: Comparison of limits on WIMPs annihilating to b quarks, using the Galaxy as a target and assuming the NFW profile, from HESS [111], ANTARES [121] and IceCube (using the Galactic Centre [120] and the Galactic Halo [122]).

the Galactic Centre is significantly larger) while the background is virtually non-existent, as there are no astrophysical processes taking place which could produce high-energy gamma rays or neutrinos. Also, the expected signal flux is much less dependent on the dark matter profile than for the Galaxy [124].

Dwarf spheroidal galaxies can be separated in two groups: the 'classical' dwarf galaxies and the 'ultrafaint' dwarf galaxies, based both on their luminosity and when they were discovered. All classical dwarf galaxies were discovered before the turn of the century, while the ultrafaint dwarf galaxies were discovered later, using more sensitive experiments. The first dwarf galaxy, Sculptor, was discovered by Harlow Shapley in 1938 [125], who at first thought that he had found a galaxy cluster. Only later he realized that the separately visible objects were stars, not galaxies, and that the distance to the system was no more than 80 kpc. In the remainder of the 20th century, another eight similar systems were found: Fornax, Draco, Ursa Minor, Leo I, Leo II, Carina, Sextans and Sagittarius, all named after the constellation they were found in.

Since 2000, the Sloan Digital Sky Survey (SDSS) [126] has been systematically observing the sky and taking spectra from possibly interesting objects. Its data has been used to search for dwarf galaxies by looking for overdensities of stars, and up until now this has almost tripled the number of known dwarf spheroidal galaxies [127]. All of these new dwarf galaxies are significantly less bright than the classical dwarf galaxies, which explains why they were not found before. The expectation is that even more dwarf galaxies will be found with the next generation of experiments. Eight candidates have already been found in data from the Dark Energy Survey [128], though further studies are necessary to determine whether these

are indeed dwarf spheroidal galaxies.

The dark matter profiles of dwarf spheroidal galaxies can be studied by observing the kinematics of stars in these galaxies. Current observations are not precise enough to determine which of the profiles discussed in section 2.4.1 gives the best description of the dark matter distribution in these objects. However, it has been shown that for dwarf spheroidal galaxies, the integrated J-factor is relatively insensitive to the choice of dark matter profile [123]. For most searches, the NFW profile is used as a benchmark for historic reasons.

All gamma-ray experiments discussed in section 2.3.3 have performed searches for gamma rays from dark matter annihilations in dwarf galaxies. At low WIMP masses, the most stringent limit comes from the Fermi-LAT satellite, which due to its large field of view has been able to observe all currently known dwarf spheroidal galaxies. Combining data from 6 years of operation and 15 dwarf spheroidal galaxies, they were able to exclude the thermal relic cross section for WIMP masses lower than approximately 10 to 100 GeV, depending on the annihilation channel [124].

Due to their smaller field of view, IACTs are only able to observe one dwarf galaxy at a time. This target needs to be carefully selected, balancing a large expected signal (high J-factor) with robustness (as especially for the ultrafaint dwarf galaxies, the uncertainty on the J-factor can be large). The current best limit comes from 160 hours of observing Segue 1 with the MAGIC telescopes [129], and as can be seen in figure 2.10 for WIMP masses above approximately 1 TeV it is better than the Fermi-LAT limit. Other interesting limits were produced by VERITAS, also studying Segue 1 [130] (an erratum was published recently [131]), and HESS, studying the Sagittarius, Coma Berenices, Fornax, Carina and Sculptor dwarf galaxies [132].

For neutrino experiments, the only published limit was produced by the IceCube experiment [133]. They studied Segue 1, Ursa Major II, Coma Berenices and Draco, and a stacking of the first two. Except for direct annihilation to neutrinos and a few other channels at high WIMP masses, the gamma-ray limits are significantly better than the neutrino limits, as can also be seen in figure 2.10.

2.4.4. Galaxies and galaxy clusters

In addition to the spheroidal dwarf galaxies discussed above, other galaxies close to the Galaxy are interesting targets for dark matter searches as well, as they are expected to contain a lot of dark matter. Unlike for the dwarf spheroidal galaxies, a gamma-ray background from cosmic-ray interactions needs to be taken into account, since these galaxies are rich in gas and dust and star formation is ongoing. The main targets are the Large and Small Magellanic Clouds (at distances of 50 and 60 kpc respectively), which are usually classified as irregular dwarf galaxies, and M31 (at a distance of about 780 kpc), which is a spiral galaxy similar to the Galaxy. These three galaxies all appear as extended sources when seen from Earth, though depending on the halo profile most of the dark matter signal will come from the inner degree. An interesting consequence of this is that with a good direction resolution, it would be possible to study the spatial profile of a possible excess and see if it is consistent with what would be expected from dark matter.

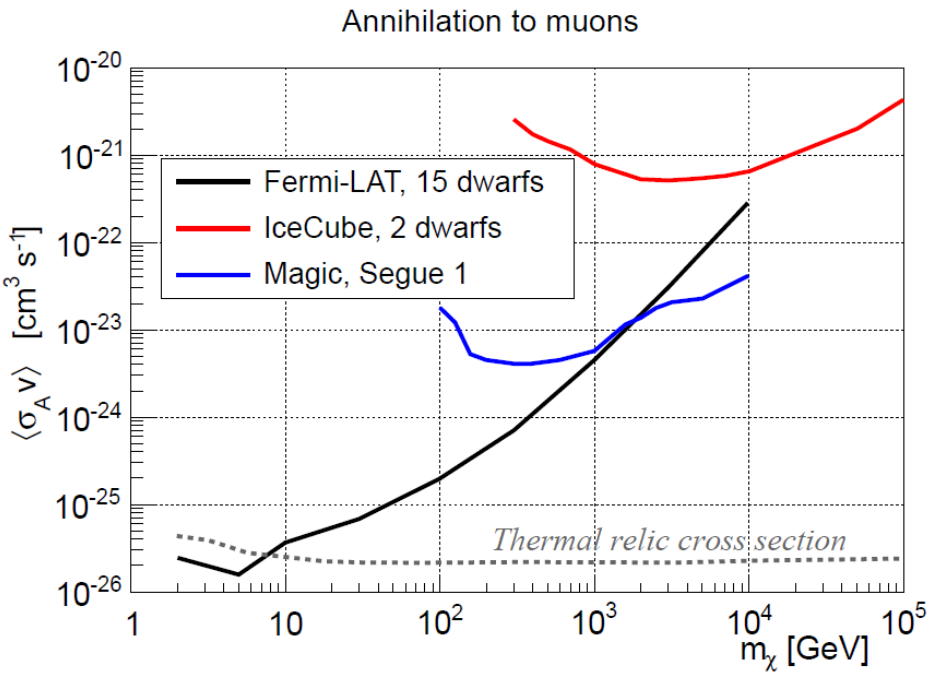


Figure 2.10.: Comparison of limits on WIMPs annihilating to muons, using dwarf spheroidal galaxies as a target and assuming the NFW profile. The limits shown here are from the Fermi-LAT combined analysis of 15 dwarf galaxies [124], observations of Segue 1 by MAGIC [129], and the IceCube combined analysis of 2 dwarf galaxies (Segue 1 and Ursa Major II) [133].

In case of the Large Magellanic Cloud (LMC), a search for annihilating dark matter has been done using data from the Fermi-LAT telescope [134]. Estimating the J-factor of the LMC is difficult because the gravitational centre and inclination are not known to great precision. However, the authors of [134] estimate that the expected flux is larger than that of the most interesting dwarf spheroidal galaxies. This advantage is partially tempered by the larger background. The gamma-ray emission detected from the LMC is compatible with the expected background from standard astrophysical sources, though there is a small excess. The limits which were set on the annihilation cross section are compared to the limits from the Fermi-LAT combined dwarf analysis and several other limits in figure 2.11. From this figure, one can see that the limits from the combined dwarf analysis are significantly better over the full WIMP mass range.

The M31 galaxy has been studied both in gamma rays (using data from Fermi-LAT) and in neutrinos (using data from IceCube). The Fermi-LAT collaboration has measured the gamma-ray spectrum from M31 and used that to conservatively constrain the annihilation cross section for a 100 GeV WIMP annihilating to b quarks, assuming no background and an Einasto profile [135]. The IceCube collaboration has set an upper limit on the annihilation cross section for different channels and WIMP masses, assuming an NFW profile [133], both without and with (an optimistic estimate for) substructure. For Fermi-LAT, the limit from the combined dwarf galaxies is better than that for M31 without substructure, as can be seen in figure 2.11. For IceCube, the limits are comparable, as can be seen in figure 2.12, which shows a comparison of limits on the annihilation cross section from neutrino experiments,

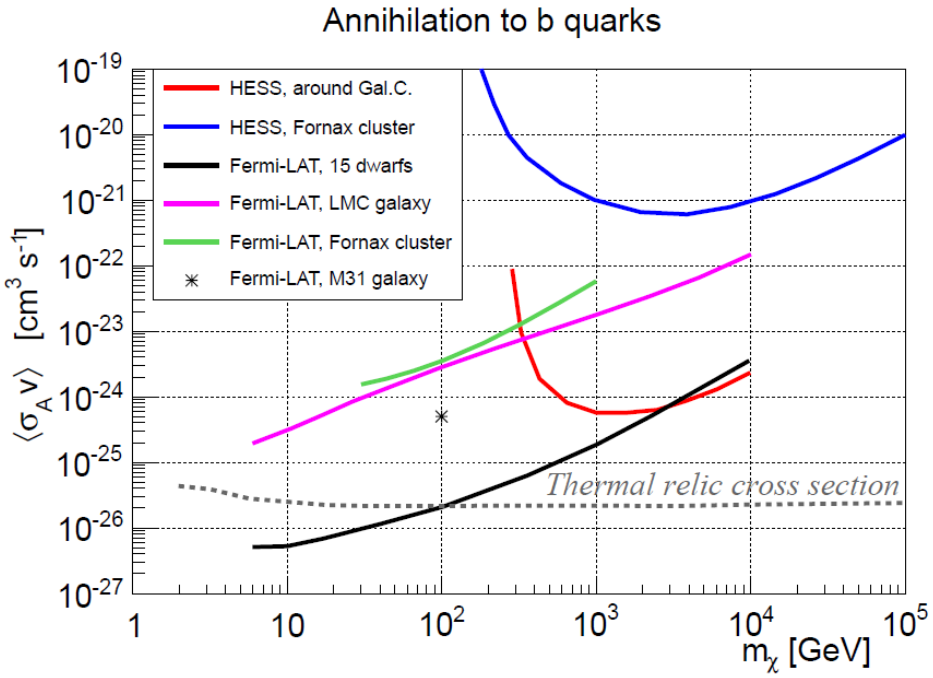


Figure 2.11.: Comparison of limits on WIMPs annihilating to b quarks from gamma-ray experiments (HESS and Fermi-LAT), assuming the NFW profile (except for the analysis of the M31 galaxy, which assumes the Einasto profile) without substructure.

using different targets.

In addition to galaxies, galaxy clusters are also considered as targets for dark matter searches. Galaxy clusters are the largest gravitationally bound structures in the Universe. Typical masses are 10^{14} to $10^{15} M_{\odot}$, which is a factor 100 to 1000 more than large galaxies like M31. As the name implies, each galaxy cluster contains a number of galaxies (both dwarf galaxies and normal galaxies). However, most of the visible matter in galaxy clusters is in the form of intracluster gas, which is visible due to it being compressed in the large potential well: the gas is heated to temperatures where it starts to emit X-ray radiation [136]. The visible matter is only 10-20% of the total mass of the galaxy cluster, and the rest is assumed to be made up of dark matter, which makes galaxy clusters interesting targets for dark matter searches.

For gamma-ray searches, there are large (expected or already detected) backgrounds in galaxy clusters, for example from cosmic-ray interactions in the intracluster gas and smaller sources that are part of the cluster. Therefore, deciding which galaxy cluster is best suited for a dark matter search depends not only on the expected signal, but also on the expected backgrounds. For example the Virgo cluster, which has the highest J-factor of all galaxy clusters in the Northern Hemisphere, is problematic because it hosts M87, an Active Galactic Nucleus which is bright in gamma rays [137]. This is less problematic for neutrino experiments, because the Virgo cluster has not yet been detected in neutrinos (in fact, no high-energy neutrino point source has been discovered yet [138]). On the other hand, if a neutrino flux from a galaxy cluster is detected, it will be complicated to show that it was produced by dark

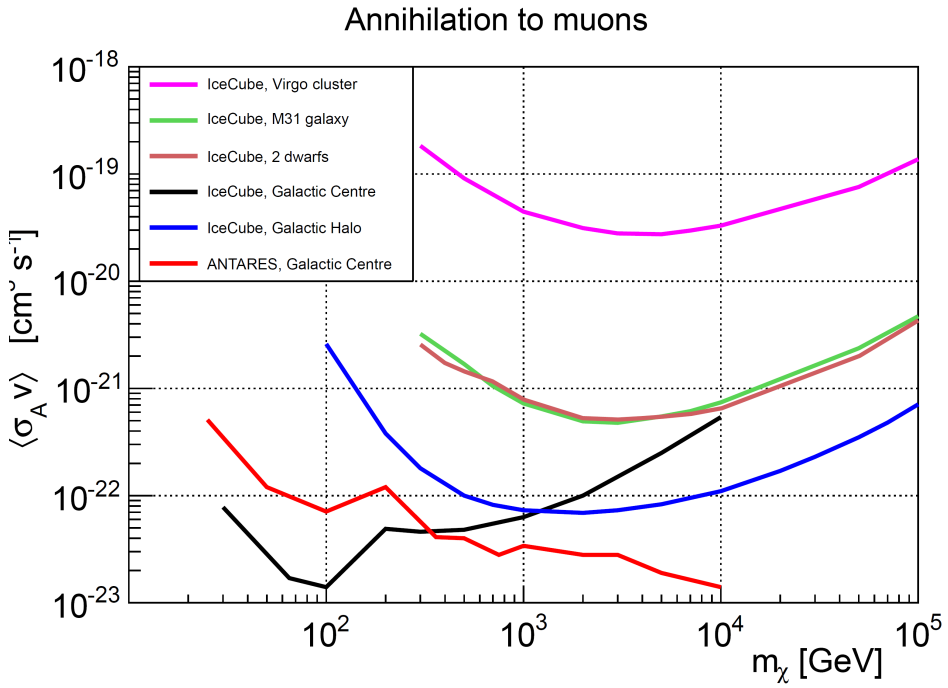


Figure 2.12.: Comparison of limits on WIMPs annihilating to muons from neutrino experiments (IceCube and ANTARES), assuming the NFW profile without substructure.

matter annihilations and not in other processes.

For neutrino experiments, the best limits come from the analysis of the Virgo cluster described in [133]. When assuming an optimistic value for the boost factor from substructure, this gives the best limit for WIMPs with high masses and annihilation to muons or taus. However, when not taking substructure into account, the result is more than an order of magnitude worse than the limit from observing the M31 galaxy or dwarf galaxies, as is shown in figure 2.12.

The most recent published analysis using Fermi-LAT data studied six clusters (Fornax, AWM 7, M49, NGC 4636, Centaurus and Coma) with the best limit coming from analyzing the Fornax cluster [139]. This analysis is from 2010, but it is likely that it will be updated soon, since a combined analysis searching for spatially extended gamma-ray emission from 50 galaxy clusters was published recently [140], and this data could also be used for a dark matter search. Limits have also been produced by the HESS collaboration studying the Fornax cluster [141] (with erratum in [142]) and the VERITAS collaboration studying the Coma cluster [143]. The limits from Fermi and HESS (both using the Fornax cluster as a target) are shown in figure 2.11, while the VERITAS result is left out since it is approximately an order of magnitude worse than the HESS result. Like for neutrinos, if substructure is not taken into account, the limits from galaxy clusters are significantly worse than those from analyses using dwarf galaxies or the Galactic centre as a target.

2.4.5. Sun and Earth

As stars and planets move through the dark matter halo of the Galaxy, WIMPs occasionally scatter off the nuclei inside these objects. As a result of these interactions, they sometimes

lose enough energy to become gravitationally bound to the object. Through further scatterings, they will then lose more energy and accumulate in the centre of the star or planet, which means that the dark matter annihilation rate there will be relatively high. This happens in every star and planet, but since the Sun and Earth are by far the closest objects, it makes most sense to consider these two targets.

The Sun and the Earth are different from other indirect detection targets, because no limits on the annihilation cross section are set. In the case of the Sun, it is assumed that equilibrium between the capture and annihilation rate has been reached [144]. This means that the flux is no longer proportional to the velocity-averaged annihilation cross section, but to the capture rate and thus the WIMP-nucleon interaction cross section. This is the same quantity that is constrained by direct detection experiments. As was already mentioned in section 2.3.1, since the Sun consists mainly of protons, limits from indirect detection are mainly competitive in constraining the spin-dependent WIMP-proton interaction cross section, and less interesting for the spin-independent part of the total cross section. These limits are complementary to the ones from direct detection experiments because the astrophysical uncertainties are different. One example is that since the annihilating WIMPs have accumulated over billions of years, the limits are much less dependent on variations in the local dark matter density over time [3]. On the other hand, the indirect detection limits do depend on the annihilation channel while the direct detection limits do not.

Since neutrinos are the only particles that can escape dense environments like stars and planets, only neutrino experiments can study WIMP annihilation in the Sun and Earth. Currently, the IceCube [145], Super-Kamiokande [146] and BAKSAN [147] collaborations have published limits on the WIMP-nucleon interaction cross section using observations of the Sun. These limits are compared to limits from direct detection experiments for a limited WIMP mass range in figures 2.13 and 2.14. From this plot one can see that as expected, the indirect detection experiments are mainly competitive for the spin-dependent WIMP-proton interaction cross section.

The Earth was created too recently to have reached equilibrium between capture and annihilation yet, so the annihilation rate is not proportional to the WIMP-nucleon interaction cross section [144]. Since the total amount of dark matter and its distribution in the centre of the Earth are not known, it is also not possible to determine the proportionality factor between the annihilation rate and the velocity-averaged annihilation cross section. For this reason, searches looking for WIMP annihilations in the centre of the Earth typically put limits on the annihilation rate or the muon flux coming from the centre of the Earth. The latest available limits are from the AMANDA [148] and Super-Kamiokande [149] experiments. No significant flux of neutrinos was found, and the limits on the muon flux from the centre of the Earth are on the order of 10^3 km^{-2} y^{-1}.

2.5. Neutrinos from dark matter annihilations in dwarf galaxies, the M31 galaxy and the Virgo cluster

The subject of the remainder of this thesis is a search for neutrinos produced in dark matter annihilations in five dwarf galaxies in the Northern Hemisphere, the M31 galaxy and the Virgo cluster, using three years of data from the completed IceCube detector. In this section,

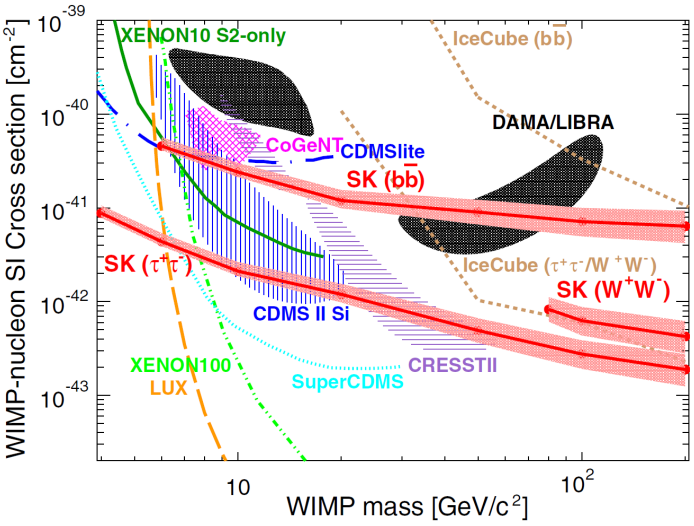


Figure 2.13.: The 90% C.L. exclusion limits (lines) and signal regions (shaded areas) for the spin-independent WIMP-nucleon interaction cross section for a number of experiments, comparing results from direct detection experiments (DAMA/LIBRA [40], CoGeNT [41], CRESST II [42], CDMS II Si [43], SuperCDMS [39], XENON10 S2-only [48], XENON100 [37], LUX [38]) and indirect detection experiments (IceCube [145] and SuperKamiokande [146]). Taken from [146].

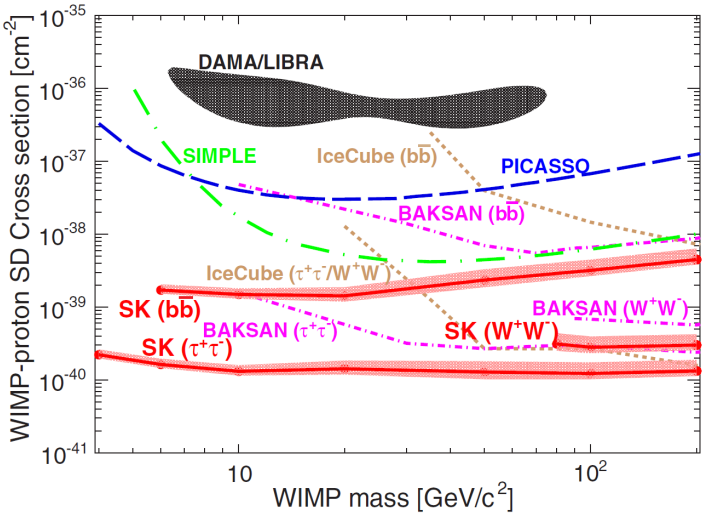


Figure 2.14.: The 90% C.L. exclusion limits (lines) and signal regions (shaded areas) for the spin-dependent WIMP-nucleon interaction cross section for a number of experiments, comparing results from direct detection experiments (PICASSO, SIMPLE, DAMA/LIBRA [40]) and indirect detection experiments (IceCube [145], BAKSAN [147] and SuperKamiokande [146]). Taken from [146].

the properties of the expected signal flux will be discussed.

The expected flux is given by equation 2.3, which depends on the velocity-averaged annihilation cross section (the quantity to be measured or constrained), the dark matter mass, the energy spectrum (which depends on the annihilation channel) and the integrated J-factor (which depends on the halo profile, distance to the target and mass of the target). To cover a large range of dark matter candidates, five benchmark channels (annihilation to muons, taus, W bosons, b quarks and neutrinos) and 19 masses ranging from 30 GeV to 100 TeV are considered. For each combination of a benchmark channel and a WIMP mass, an excess will be searched for and if this is not found, a limit will be put on the annihilation cross section for annihilation to this channel for this WIMP mass.

The targets being considered are extragalactic targets in the Northern Hemisphere, as for the IceCube detector the background in the Northern Hemisphere is much lower than in the Southern Hemisphere (see chapter 3). In each category (dwarf galaxies, galaxies and galaxy clusters) the targets with the highest J-factors will be considered. For galaxy clusters and galaxies, the Virgo cluster and the M31 galaxy have a significantly higher J-factor than any of the other candidates. For dwarf galaxies, on the other hand, there are a number of targets with J-factors in the same range, so the five with the highest J-factors (according to [127]) are considered. Since these are similar targets, it is also possible to combine data from all five to get a better limit.

For all targets, an NFW profile will be assumed, to enable comparisons with other experiments. As was discussed in section 2.4, it is currently not clear which profile best describes the actual dark matter distribution, and the profile could be different for haloes of different sizes. However, with the current direction resolution of IceCube it is not possible to study the possible extension of the signal for any of the considered profiles (if substructure is ignored), so one can simply scale the limit for the NFW profile up or down to get the limit for a different profile. Since the value of the boost factor and the expected change in the spatial profile from subclustering are still uncertain, substructure will be ignored and a smooth NFW profile will be assumed for each target.

Table 2.1 contains the locations, distances and J-factors (assuming a smooth NFW profile) of each of the seven targets that are considered. This shows that the J-factors of the Virgo cluster and the M31 galaxy are smaller than those of the best dwarf galaxies, which will lead to a worse sensitivity to the annihilation cross section. However, for these larger targets there could be a significant boost factor from substructure. For this reason, the Virgo cluster and the M31 galaxy are still interesting targets for discovery, though less so for setting limits. Since substructure is expected to make the profile more extended, an extended signal from the M31 galaxy and the Virgo cluster will also be considered.

The energy spectrum of the signal neutrino flux depends strongly on the annihilation channel. For direct annihilation to neutrinos (which is assumed to be $\frac{1}{3}(\nu_e\bar{\nu}_e + \nu_\mu\bar{\nu}_\mu + \nu_\tau\bar{\nu}_\tau)$), the energy spectrum will be a delta function at the WIMP mass, since typical WIMP speeds are negligible. For the other annihilation channels, the neutrino energy never exceeds the WIMP mass, but the spectra differ significantly. This can be seen in figure 2.15, where muon neutrino energy spectra for the different annihilation channels considered in this analysis are shown. The differences in the spectra can be understood as follows: W bosons, taus and

Target	Type	RA [deg]	Dec [deg]	Dist. [kpc]	$\log_{10}(J$ / GeV^2 cm$^{-5})$
Segue 1	Dwarf	151.767	16.082	23	19.5 ± 0.29
Ursa Major II	Dwarf	132.875	61.310	32	19.3 ± 0.28
Willman 1	Dwarf	162.343	51.051	38	19.1 ± 0.31
Coma Berenices	Dwarf	186.746	23.919	44	19.0 ± 0.25
Draco	Dwarf	260.052	57.915	76	18.8 ± 0.16
M31	Galaxy	10.685	41.269	785	19.2 ± 0.1
Virgo	Cluster	187.704	12.391	16800	18.5

Table 2.1.: The list of targets considered in the analysis described in this thesis, including their locations [150], their distances and their J-factors assuming a smooth NFW dark matter profile [127, 151, 152]. For the Virgo cluster, no uncertainty on the J-factor is available.

muons can directly decay to neutrinos, so the energy spectrum will be relatively hard. On the other hand, when dark matter annihilates to b quarks, neutrinos are only produced after hadronization, when the parent particles have lost much of their energy already. For this reason, the spectrum for annihilation to b quarks is softer than for the other channels.

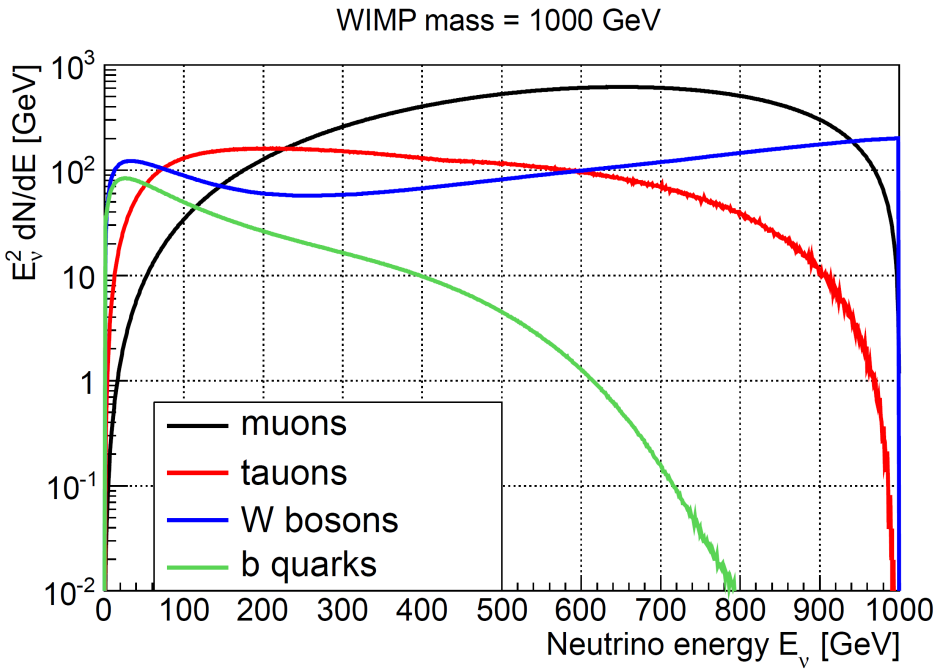


Figure 2.15.: Comparison of muon neutrino energy spectra at the source multiplied by E_ν^2 for annihilation to different benchmark channels, for a 1 TeV WIMP. This plot is based on simulations of annihilating WIMPs which were done with Pythia 8.1 [153].

2.5.1. Neutrino oscillations

Figure 2.15 shows the spectra for muon neutrinos, which is the neutrino flavour that is considered in the analysis described in this thesis (the reason for this will be explained in chapter 3). However, while travelling from the source to the Earth neutrinos undergo flavour oscillations, meaning that for example a neutrino which was produced as a tau neutrino can be detected at the Earth as a muon neutrino. This is due to the fact that for neutrinos the

interaction (flavour) eigenstates $\nu_{e,\mu,\tau}$ are different from the propagation (mass) eigenstates $\nu_{1,2,3}$, and the mass differences between the mass eigenstates are small [154]. Mass and flavour eigenstates are related by [154]

$$\nu_\alpha = \sum_{i=1,2,3} U^*_{\alpha i}\nu_i \tag{2.10}$$

with $U_{\alpha i}$ the elements of the matrix U which describes the mixing. The elements of this matrix depend solely on three neutrino mixing angles and a complex phase, and there are a number of experiments currently active whose goal is to measure these constants to high precision.

The probability to detect a neutrino with energy E as flavour eigenstate β if it was produced as flavour eigenstate α and has travelled a length L is given by [154]

$$\begin{aligned} P_{\nu_\alpha\to\nu_\beta}(L,E) = \delta_{\alpha\beta} \quad &- \quad 4\sum_{k>j} Re[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}]\sin^2\left(\frac{\Delta m^2_{kj}L}{2E}\right) \\ &+ \quad 2\sum_{k>j} Im[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}]\sin\left(\frac{\Delta m^2_{kj}L}{2E}\right) \end{aligned} \tag{2.11}$$

where j and k run over the neutrino mass eigenstates and Δm^2_{kj} is the mass difference between mass eigenstates j and k. In case of an extended source, where there is an uncertainty on L, one should integrate over the size of the source. For the analysis described in this thesis, the extension of the source (which is approximately 50 pc for dwarf galaxies and larger for the M31 galaxy and Virgo cluster) is much larger than the so-called oscillation length, which is defined as

$$L_{osc} = \frac{2E}{\Delta m^2_{kj}}. \tag{2.12}$$

and is $\lesssim 2.5\,\mu$pc for the neutrino energies relevant for this analysis. That means that integrating over the extension of the source is equivalent to taking the average of equation 2.12. This results in the following average oscillation probability:

$$P_{\nu_\alpha\to\nu_\beta} = \delta_{\alpha\beta} - 2\sum_{k>j} Re[U^*_{\alpha k}U_{\beta k}U_{\alpha j}U^*_{\beta j}]. \tag{2.13}$$

which no longer depends on L or E. Using the current world average values for the neutrino mixing angles as found in [11], one can calculate that the probability to be detected as a muon neutrino on Earth is approximately 25% for a neutrino which is produced as an electron neutrino, 37% for a neutrino which is produced as a muon neutrino and 38% for a neutrino which is produced as a tau neutrino. Using these values, one can calculate the muon neutrino flux at Earth from the signal fluxes at the source for all three flavours.

3. Detecting neutrinos from dark matter annihilations

If there is indeed a neutrino flux produced by annihilating dark matter in nearby galaxies and galaxy clusters (as described in section 2.5), it might be possible to detect this flux with large-scale neutrino detectors. Usually, these are called neutrino telescopes because their main goal is to detect astrophysical neutrinos and study their sources. Currently, there are three active neutrino telescopes: NT-200 (often called 'the Baikal underwater neutrino telescope') [155], ANTARES [156] and IceCube [157]. All three use a natural medium (lake water, sea water or ice) as detection medium, are built at large depths and study neutrinos created in natural processes. Examples of these are solar neutrinos, atmospheric neutrinos, neutrinos produced in Active Galactic Nuclei or Gamma Ray Bursts, and neutrinos produced in dark matter annihilations.

The subject of this thesis is an analysis of IceCube data. However, before IceCube is described in section 3.3, sections 3.1 and 3.2 discuss subjects that are relevant to all neutrino telescopes: the detection principle and the main backgrounds for a search for neutrinos from dark matter annihilations. In the final section of this chapter, the Monte Carlo simulation of both signal and background is discussed.

3.1. Detection principle

As neutrinos only interact with other particles via the weak force, the probability of a neutrino interacting with another particle is extremely small. The interaction cross section increases with neutrino energy [158], but even at 100 TeV (the highest energy relevant for this analysis) the probability of a neutrino interacting when it passes through a meter of water is of the order of 10^{-8}. As a consequence of this, one needs very large detectors to be able to detect neutrino fluxes of the order of what is expected from dark matter annihilations and other astrophysical processes. This is why neutrino telescopes have been built in already existing bodies of water or ice like the Mediterranean Sea or the South Pole ice.

When a neutrino interacts with another particle, high-energy charged particles like electrons or muons can be created, which produce Cherenkov radiation while passing through a medium. Detecting this radiation is the basic detection principle of neutrino telescopes, and this section will start with a discussion of Cherenkov radiation.

3.1.1. Cherenkov radiation

If the speed of a charged particle in a medium v is larger than the speed of light in that medium (which is given by $c_m = c/n$, with n the refraction index of the medium), it will start to emit coherent electromagnetic radiation, called Cherenkov radiation [159]. This is due to the fact that the charged particle polarizes the matter close to it, and for $v > c_m$ this polarization adds coherently. Using that according to special relativity, the energy E and

mass m of a particle are related by $E = mc^2/\sqrt{1 - v^2/c^2}$, one can calculate that Cherenkov radiation is emitted when the energy of a charged particle is larger than

$$E_{min} = \frac{m}{\sqrt{1 - \frac{1}{n^2}}}. \tag{3.1}$$

In most cases, the products of the interactions described in the next section will have energies far exceeding their rest mass, so they will indeed emit Cherenkov radiation.

Cherenkov radiation is emitted in a cone around the track of the charged particle, with an angle of emission given by

$$\theta_c = \cos^{-1}\left(\frac{1}{\beta n}\right), \tag{3.2}$$

so for very relativistic particles (with $\beta \approx 1$) the angle is independent of the energy of the particle and only depends on n. The spectrum of the Cherenkov photons is given by [160]

$$\frac{d^2N}{dxd\lambda} = \frac{2\pi\alpha}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2}\right) \tag{3.3}$$

with α the fine-structure constant, λ the wavelength of the photons and x the unit track length. For both ice and water, n is approximately 1.35 in the wavelength range from 300 to 500 nm, which is the relevant range for the light sensors used in current experiments. That means that one would expect approximately 270 photons in this wavelength range per centimeter of track length.

3.1.2. Neutrino interactions with matter

It is complicated to describe neutrino-nucleon interactions at energies below about 20 GeV, as there are several different interactions that can take place. However, once the neutrino energy is larger than that (which is the case for most of the neutrinos detected in IceCube) the dominant type of interaction is a weak interaction between a quark and a neutrino, which is called Deep Inelastic Scattering (DIS) [161]. This interaction is dominant because at these energies, neutrinos can resolve the structure of a proton or neutron. As in all weak interactions, either a W boson (in a charged-current interaction, abbreviated as CC interaction) or a Z boson (in a neutral-current interaction, abbreviated as NC interaction) is exchanged. The Feynman diagrams for these interactions are shown in figure 3.1. In case of a CC interaction, a charged lepton of the same flavour as the neutrino is produced and the quark flavour changes. In a NC interaction the interaction particles do not change, only energy is transferred.

Both types of interaction will cause a hadronic shower with a length of about 5 to 10 meters (see figure 3.2) due to the nuclear recoil caused by the energy transfer from the neutrino to the quark. In this hadronic shower, many types of charged particles are created which can produce Cherenkov light if their energies are larger than E_{min} (defined in equation 3.1). In NC interactions, the hadronic shower is the only visible effect of the interaction, while in CC interactions the outcoming charged lepton produces additional Cherenkov light.

In ν_e CC interactions, the high-energy electron or positron which is created will produce an electromagnetic shower: the initial electron or positron produces bremsstrahlung pho-

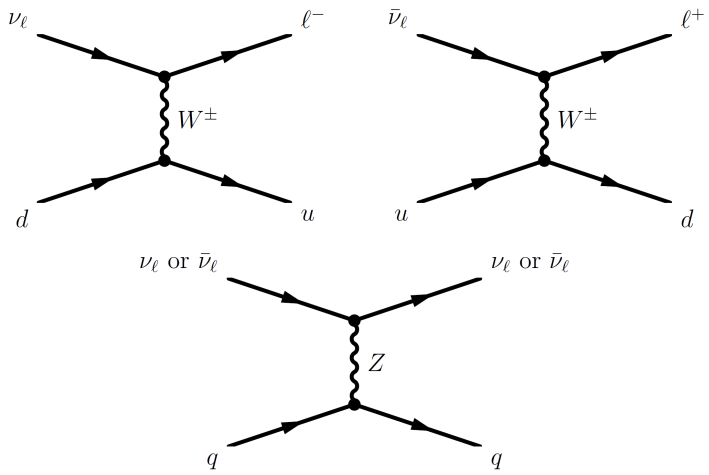


Figure 3.1.: Feynman diagrams for the possible interactions between a quark and a neutrino. The upper two diagrams show charged-current interactions, where a W boson is exchanged and a charged lepton is created. The lower diagram shows a neutral-current interaction, where a Z boson is exchanged and only the energy of the interacting particles changes.

tons, which will produce electron-positron pairs, which will produce bremsstrahlung photons, etcetera. This continues until the energies of the particles in the shower are too low to produce new particles. The total length of the tracks of all charged particles is directly proportional to the energy of the initial electron or positron, with a proportionality constant of 5.31 m/GeV [162]. As the track length in its turn is proportional to the number of Cherenkov photons (see equation 3.3), one can calculate the initial energy if the amount of Cherenkov light is known.

As can be seen in figure 3.2, below approximately 10 TeV an average electromagnetic shower is slightly less elongated than an average hadronic shower. Also, the total Cherenkov light output is higher for electromagnetic showers, as hadronic showers contain more neutral particles and the Cherenkov threshold for charged hadrons is larger than for electrons and positrons. Since there are more different processes taking place in hadronic showers than in electromagnetic showers, the fluctuations in their light output and elongation are also larger. The higher the energy, the more similar hadronic showers and electromagnetic showers become, as the fraction of the hadronic shower which is electromagnetic (originating from neutral pions decaying to photons) increases [163].

In ν_μ CC interactions a muon is produced, with a range significantly longer than that of a shower for the energy range shown in figure 3.2. In section 3.1.3, the dependence of the muon range on the muon energy will be discussed. As a result of the long range, the direction of a muon can be reconstructed more easily than that of a shower. The muon direction is related to the neutrino direction and the average angle between the two directions can be approximated by [165]

$$\langle \theta_{\mu-\nu} \rangle \simeq \frac{0.7^\circ}{(E_\nu / 1 \text{ TeV})^{0.7}}, \tag{3.4}$$

so for energies above 1 TeV the muon and neutrino are almost parallel. The result is that, as will be shown in section 4.1, the direction of the neutrino can be reconstructed relatively well. This is very useful when searching for neutrino sources. Also, for neutrino telescopes

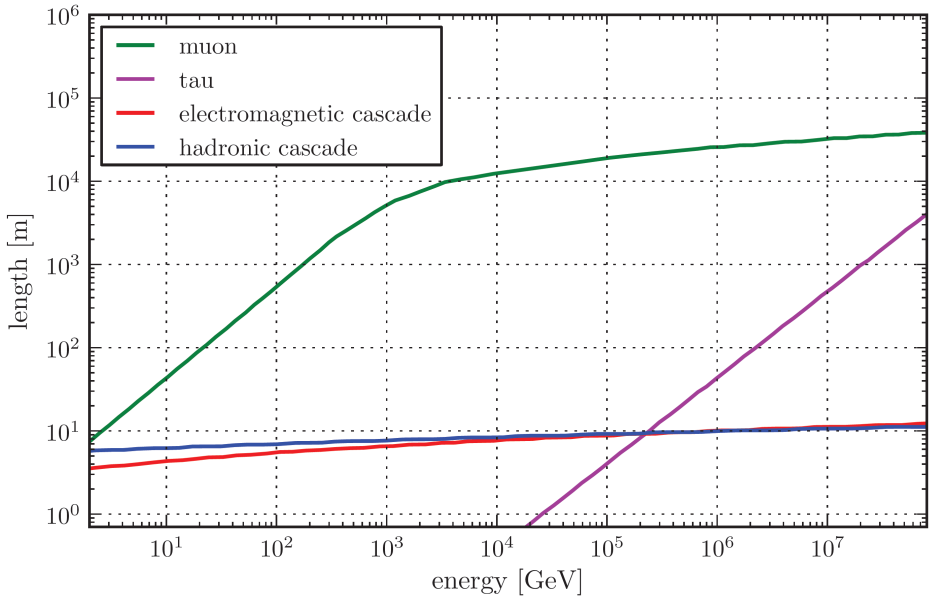


Figure 3.2.: Range of the four possible products of neutrino-quark interactions, as a function of their energy. Taken from [164].

in a natural environment the effective detector volume is increased for ν_μ CC interactions: even if an interaction takes place far outside the detector, it is possible to detect the result as long as the initial muon energy is high enough.

The final event type are ν_τ CC interactions, which produce a tau. For PeV energies and higher, the distance between the production and decay positions of the tau can be large enough that the separation is visible. However, in the energy range relevant for this analysis (up to 100 TeV), the separation is at most a few meters, as can be seen in figure 3.2. In most cases, the tau decay creates a second (hadronic or electromagnetic) shower, but in approximately 17% of the cases the tau decays to a muon. This gives a signature similar to ν_μ CC interactions, but with a lower-energy muon.

For the neutrino telescopes discussed here it is often not possible to determine which type of neutrino interaction caused a certain event. However, it is possible to divide neutrino events in two types: tracks (events with a clearly visible track, almost always caused by a muon) and cascades (all other events). As direction resolution is very important for the analysis described in this thesis, only track events will be used and cascade events will be removed from the data sample as much as possible.

3.1.3. Muon energy loss

The energy loss from radiating off Cherenkov photons is only a very small part of the total energy loss of charged particles while moving through ice or water. In this section, muon energy losses will be discussed, since tracks are the detection channel used in the analysis described in this thesis. The four main processes for muon energy loss in ice or water are ionization, bremsstrahlung, pair production and photo-nuclear interactions [11]. The first of these is a continuous process, while the latter three (which are usually collectively called

'radiative processes') are stochastic in nature.

The average muon energy loss in a material can be written as

$$-\left\langle \frac{dE_\mu}{dx} \right\rangle = a(E_\mu) + b(E_\mu) \cdot E_\mu, \tag{3.5}$$

where $a(E_\mu)$ describes the energy losses due to ionization (which depend logarithmically on the muon energy [166]) and $b(E_\mu) \cdot E_\mu$ describes the energy losses due to radiative processes, which are approximately proportional to the muon energy. Assuming that a and b are constant (as the energy dependence is known to be small), the muon range R_μ can be determined by integrating equation 3.5, and the result is

$$R_\mu = \frac{1}{b} \ln\left(1 + \frac{b}{a} E_0\right), \tag{3.6}$$

with E_0 the initial energy of the muon.

For ice, a is approximately 0.26 GeV/m and b is approximately $0.36{\cdot}10^{-3}$ m^{-1} [163], while for water these values are 0.27 GeV/m and $0.35{\cdot}10^{-3}$ m^{-1} [167]. From this, one can calculate that ionization is dominant below approximately 700 GeV for ice and 800 GeV for water, while radiative processes (dominated by pair production) are dominant for higher energies. This is also clearly visible in figure 3.2: below 500 GeV, the range grows linearly with the initial energy (since if $b \cdot E_0 \ll a$, equation 3.6 simplifies to $R_\mu \approx E_0/a$), while above a few TeV, the range grows logarithmically (since if $b \cdot E_0 \gg a$, equation 3.6 simplifies to $R_\mu \approx \ln(E_0 \cdot b/a)/b$),

3.2. Backgrounds

Like many physics experiments, neutrino telescopes do not only detect signal events but also background events, and the background rate is much larger than the expected signal rate. What is considered background and what is considered signal depends on the specific analysis, but in most cases the signal consists of neutrinos produced far away from Earth (for example in dark matter annihilations). In that case, the main backgrounds are atmospheric muons and atmospheric neutrinos, which will be discussed in this section.

Both atmospheric muons and atmospheric neutrinos are produced when cosmic rays hit the atmosphere of the Earth. The cosmic ray flux has been measured over twelve orders of magnitude (from MeV to EeV energies) and is known to be approximately isotropic, to mainly consist of protons and other atomic nuclei, and to have a steeply falling spectrum [11]. The most precise method to study cosmic rays is to detect them before they enter the atmosphere, and this is done by satellite and balloon experiments. However, at energies above approximately 1 PeV, the flux is too low to do this. The solution are ground-based experiments, which can measure the high-energy component indirectly by detecting secondary particles created in the interaction with the atmosphere. Unfortunately, measurements of the energy and type of the primary particles are less precise when using this method. The measured spectrum from 10 TeV per nucleon upwards, multiplied by $E^{2.6}$ to increase the visibility of structures in the spectrum, is shown in figure 3.3.

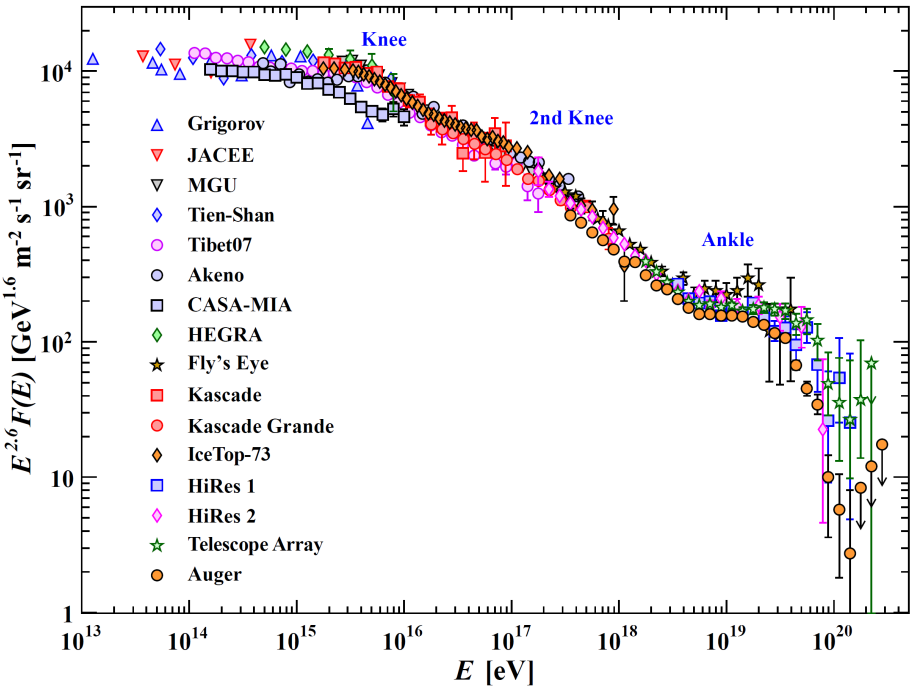


Figure 3.3.: The cosmic ray spectrum as measured by air-shower experiments (except for Grigorov, which was located on a satellite, and JACEE, which was a balloon experiment), as a function of the energy per nucleus of the primary cosmic ray. The features denoted as the 'knee', the 'second knee' and the 'ankle' are the places where the index of the power law describing the spectrum changes. Taken from [11].

The first interaction of a cosmic ray typically takes place at a height of tens of kilometers. It initiates a shower of particles which keep interacting and multiplying until their energies are too low to continue to do this. Typically, the total shower will contain millions of particles. An example of a small part of such a shower is shown in figure 3.4. In the part shown here, charged pions are produced which then decay to the atmospheric muons and neutrinos seen in neutrino telescopes. The same can happen with charged kaons or heavier mesons. Both the pions and the kaons often interact with air molecules before decaying, resulting in energy loss. This does not happen to the heavier mesons, which decay more promptly. Hence, the atmospheric muons and neutrinos created in heavy meson decays are called 'prompt' while their counterparts from pion and kaon decays are called 'conventional'. The result of this is that the prompt spectrum follows that of the primary cosmic rays, while the conventional spectrum is approximately one order of magnitude steeper. As pions and kaons are produced much more often than heavier mesons, the prompt component is mostly negligible at the energies relevant for the analysis described in this thesis.

3.2.1. Atmospheric muons

Due to their relatively long lifetime of 2.2 μs [11], atmospheric muons with energies above a few GeV are able to travel tens of kilometers. However, when travelling through matter they will lose energy due to interactions (see section 3.1.3). As a result, only muons with energies above a few hundred GeV are able to pass through the large amounts of water or ice that are shielding all current neutrino telescopes. Even with this shield, the rate of atmospheric

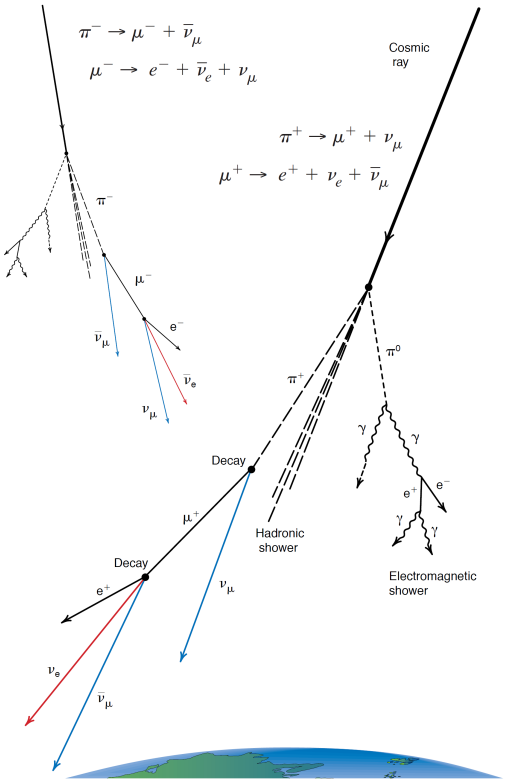


Figure 3.4.: An example of an air shower from a cosmic ray interacting in the Earth atmosphere, showing production of atmospheric neutrinos and other particles which can be detected on Earth. Taken from [168].

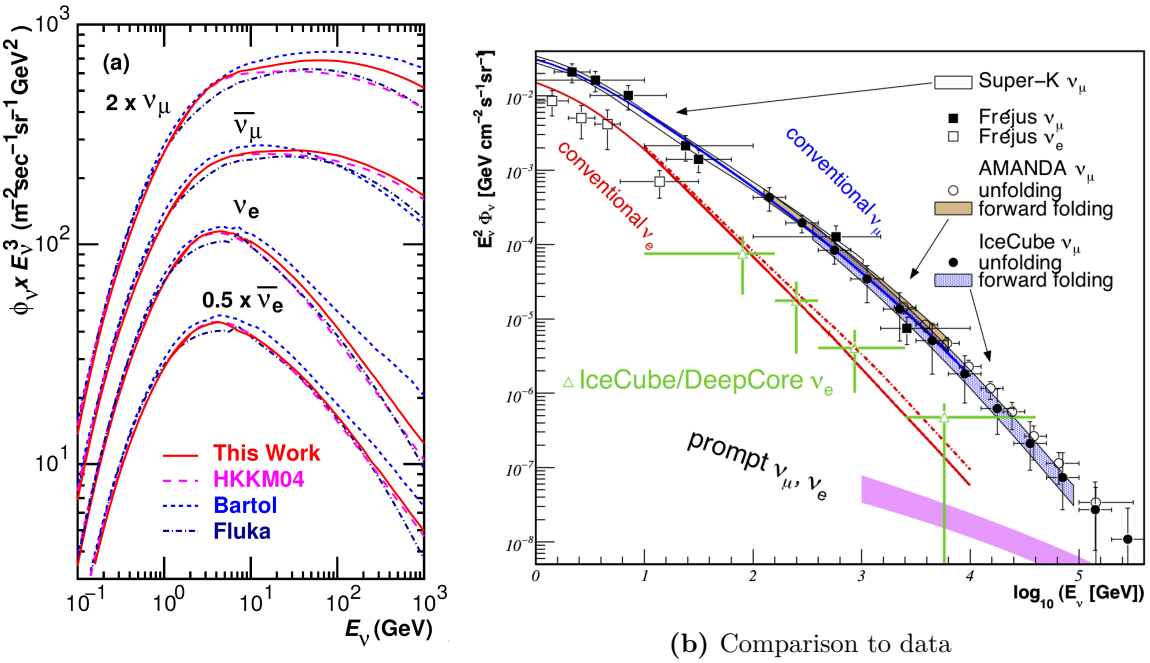
muons in neutrino telescopes is still much larger than that of (atmospheric or other) neutrinos, which makes them the main background at most energies. For example, the IceCube detector detects approximately 2500 events per second, but only once or twice per minute such an event is due to a neutrino.

There are several models available for the expected atmospheric muon flux on Earth, for example the poly-gonato model by Hörandel [169] and the H3a and H4a models by Gaisser [170]. The flux depends mainly on the composition and density of the atmosphere, the chemical composition of the primary cosmic ray flux and the model for interactions between high-energy cosmic rays and air molecules. The latter two still contain large uncertainties, due to the fact that these interactions take place at energies much higher than can be studied in a laboratory and the primary cosmic rays cannot be detected directly.

For the type of analysis described in this thesis, atmospheric muons are a reducible background. This is due to two of their properties: they are always downgoing (as they cannot pass through the Earth) and they always enter the detector from outside (unlike muons produced in neutrino interactions, as these interactions can take place inside the detector). Chapter 4 describes how to remove atmospheric muons from the sample, and the final sample used for the analysis described in chapter 5 is dominated by (atmospheric) neutrinos.

3.2.2. Atmospheric neutrinos

Currently, there are several models for the atmospheric neutrino flux, for example the Honda model [171], the Bartol model [172] and the Fluka model [173]. The resulting energy spectra for electron and muon (anti) neutrinos are shown in figure 3.5a. In figure 3.5b, two of these models are compared to a number of measurements of atmospheric neutrino energy spectra. From this one can see that the measurements of the atmospheric neutrino fluxes are not yet precise enough to determine which of the several models is correct. It is also clear that the flux of conventional atmospheric electron neutrinos is much lower than that of conventional atmospheric muon neutrinos, which is because pions and kaons mainly decay to muons and muon neutrinos. Since there are no tau neutrinos produced in pion and kaon decays, the only conventional atmospheric tau neutrinos are the result of neutrino oscillations. There is a very small prompt atmospheric tau neutrino flux, but this is suppressed by a factor of 20 compared to the other two flavours [174].



(a) Model comparison

(b) Comparison to data

Figure 3.5.: Atmospheric neutrino energy spectra. (a) compares different model predictions for the energy spectrum of atmospheric neutrinos. HKKM04 is an earlier version of the model from Honda et al, while 'This work' refers to the updated version, which is used for the simulation described in section 3.4.2. Taken from [171]. (b) shows measurements of the different atmospheric neutrino fluxes by a number of experiments, compared to the Honda model (blue and red solid lines), the Bartol model (red dotted line) and a model for the prompt neutrino flux (pink band) [174]. Taken from [175].

The conventional neutrino flux is not isotropic, but peaks at the horizon. This is because for vertical showers, the probability of a pion or kaon interacting and losing energy before it decays is larger than for horizontal showers, since the shower reaches the denser parts of the atmosphere sooner. This does not affect the prompt neutrino flux, which follows the isotropic cosmic ray flux.

On an event-by-event basis, it is impossible to tell whether a neutrino originated in a cosmic ray shower or a dark matter annihilation. This means that most of the atmospheric neutrinos cannot be removed from the event sample without also removing a large fraction of the (possible) neutrinos from dark matter annihilations. However, there are differences in the spatial (neutrinos from dark matter annihilations are more likely to come from targets which contain a lot of dark matter) and energy distributions which can be used to determine whether the final level sample contains any signal events.

3.3. The IceCube Neutrino Observatory

The past two sections discussed topics that are relevant for all neutrino telescopes, but the topic of this section is the IceCube Neutrino Observatory [157] (generally referred to as 'IceCube'), which is the detector used for the analysis described in this thesis. It is the only currently active neutrino telescope that uses ice instead of water as detection medium, and it is located at the geographic South Pole. It is also significantly larger than the other neutrino telescopes, which means that in most cases it is the most sensitive neutrino telescope. However, there are exceptions, for example certain analyses looking at targets in the Southern Hemisphere, since both ANTARES and NT-200 are located in the Northern Hemisphere.

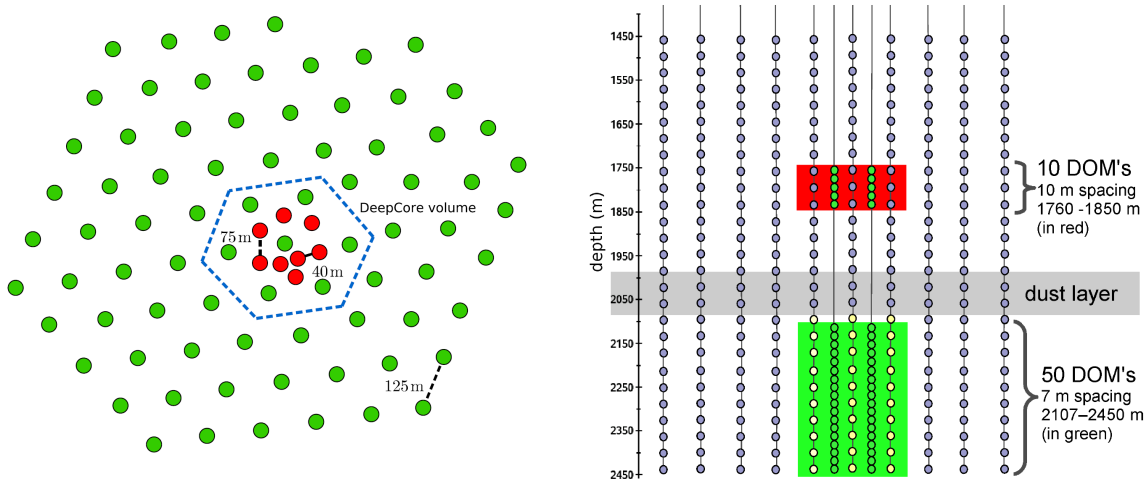
3.3.1. Layout

IceCube has an instrumented volume of approximately one cubic kilometer, and consists of 86 strings. Each string contains 60 Digital Optical Modules (DOMs) at depths between 1450 and 2450 meters. Construction started in 2004 and was finalized in 2011, deploying up to 20 strings per summer season by using hot water to drill holes into the ice, lowering the strings into these holes and letting them refreeze. Data was already taken with incomplete configurations, which are referred to as ICx, where x is the number of strings that was functional during that time. For example, the previous search for neutrinos from WIMPs annihilating in (dwarf) galaxies and galaxy clusters was done using data from IC59. For each configuration, the livetime was approximately a year, and the switch to a new configuration was made in May. The current configuration is referred to as IC86, and since there are multiple years of data taken with this configuration, usually these are denoted by the year in which they started. In the analysis described in this thesis, three years of data from the complete detector are used: IC86-2011, IC86-2012 and IC86-2013.

The IceCube strings are arranged in a hexagonal pattern with inter-string distances of approximately 125 meters, as can be seen in figure 3.6a. In both subfigures of figure 3.6, one can also see the denser sub-detector DeepCore [176], consisting of eight extra strings which are placed in between the normal pattern. In general, the IceCube strings in between and around the extra strings are also considered to be part of DeepCore. The DOMs on the eight infill strings are arranged differently from those on the 78 other strings: instead of having one DOM every 17 meters between 1450 and 2450 meters depth, the infill strings have 10 DOMs with a spacing of 10 meters between 1760 and 1850 meters and 50 DOMs with a spacing of 7 meters between 2107 and 2450 meters. The large gap between the 10th and 11th DOM on the infill strings is to avoid the so-called 'dust layer', which will be discussed in more detail in section 3.3.4. The DeepCore detector was added to lower the energy threshold of the detector: for the main array, the energy threshold is approximately 50-100 GeV, while

for DeepCore it is approximately 10 GeV [176].

On top of the ice, there is also a cosmic-ray detector called IceTop, which consists of 162 large tanks filled with ice. IceTop has been designed to detect the Cherenkov radiation from charged particles in cosmic-ray showers. In some analyses, data from IceCube and IceTop is combined, but for the analysis described in this thesis, data from IceTop was not used.



(a) Top view **(b)** Side view

Figure 3.6.: Top and side views of the IceCube detector, including the DeepCore sub-detector. The eight infill strings are shown in red in (a), and with green DOMs in (b). Based on figures taken from [177].

IceCube uses a coordinate system with the origin close to the geometric center of the complete detector (approximately 1948 meters below the surface) and the z axis pointing upwards. The y axis points towards Greenwich in the UK and is thus aligned with the Prime Meridian, and the x axis points 90 degrees clockwards from the y axis. Directions are defined by zenith and azimuth values: the zenith is the angle with the negative z axis, and the azimuth would be the angle with the positive x axis if the event would have a zenith of 90 degrees. This means that upgoing events have a zenith larger than 90 degrees. Also, the zenith of a target (for example, the ones in table 2.1) can be calculated by adding 90 degrees to the declination.

3.3.2. Hardware and readout

Each DOM consists of a glass pressure sphere filled with a downward-pointing 25 cm Hamamatsu PhotoMultiplier Tube (PMT), a grid made of mu-metal (an alloy of mainly nickel, iron and molybdenum) which cancels out the Earth magnetic field, readout electronics, a power supply and LEDs used for calibration of the detector. DOMs are attached to strings, which contain cables that are used for power and communication, both with the surface and between DOMs on the same string.

The PMT contains a photocathode which will emit an electron when a photon (for example a Cherenkov photon) is absorbed. This photo-electron is then multiplied using a series of dynode plates at increasing voltages, and in the end the signal (typically of the order of 10^7

electrons) is read out as a change in the voltage on the anode of the PMT.

The quantum efficiency (QE) of a PMT refers to the probability that an incident photon will create a photo-electron and depends on the wavelength of the photon. The PMTs used in IceCube are sensitive to wavelengths between 300 and 600 nm, and the sensitivity peak is at 420 nm, where the QE is approximately 25% [178]. The DOMs on the DeepCore infill strings contain different PMTs, which have a 35% higher QE than the normal PMTs [176]. Another important characteristic of PMTs is the noise rate, which is approximately 500 Hz for IceCube DOMs and 650 Hz for DeepCore DOMs [176, 178]. This noise is mainly due to thermal background and radioactive decays in the PMT.

The voltage at the anode of the PMT is recorded by two types of digitizers. Each DOM contains two Analog Transient Waveform Digitizers (ATWDs), which sample at a rate of 300 MHz (meaning that they have a resolution of 3.3 ns). They can only do so for approximately 425 ns, and after that, the ATWD has 29 μs deadtime while it digitizes the signal. To be able to capture longer signals, each DOM also contains a Fast Analog to Digital Converter (FADC), which records the waveform constantly with a sample rate of 40 MHz (a resolution of 25 ns). Combining the FADC and ATWDs allows the DOM to capture signals of different lengths accurately [179].

The integrated deposited charge on the anode is usually given in units of PE, which is the most likely deposited charge from a single photon (approximately 1.6 picocoulomb for a typical IceCube PMT [180]). Once the FADC detects a signal of more than 0.25 PE, it will send data to the data acquisition system at the surface. In most cases, this data will only encompass the three FADC samples with the highest amplitude. However, if one of the neighbouring or next-to-neighbouring DOMs on the same string also detects a signal of more than 0.25 PE within 1 μs, which makes it less likely that the first signal was due to PMT noise, the full FADC and ATWD waveforms are sent to the surface. This condition, which is not determined on the surface but through direct communication between the DOMs, is called Hard Local Coincidence (HLC).

As the IceCube detector is frozen in, it is impossible to replace DOMs which are no longer producing good data, for example due to communication problems (both with the surface and with neighbouring DOMs) and power-up failures. Instead, these DOMs are marked as 'bad' and any data they send to the surface is not used for reconstructions and processing. In most cases, bad DOMs have been bad since they were deployed in the ice, though sometimes issues arise later. In the dataset used for this analysis, approximately 100 DOMs (less than 2% of the total) were considered bad, which does not significantly affect the performance of the detector.

3.3.3. Triggering and filtering

The rate of hits in IceCube is high, but most are due to PMT noise and dim atmospheric muons. To select the most interesting data, a number of triggers are used. The simplest of these are based on the number of HLC hits, for example the Simple Multiplicity Trigger with multiplicity 8 (SMT8) requires at least 8 HLC hits in a time period of 5 μs. There is a similar trigger specifically for low-energy events (events below approximately 100 GeV): the SMT3 trigger requires 3 HLC hits in a time period of 2.5 μs inside the DeepCore fiducial volume.

This volume is the area surrounded by the blue dotted line in figure 3.6a and the green area in figure 3.6b.

In addition to the two SMT triggers, there are two other triggers used for the analysis in this thesis. The first is called the 'string trigger' and requires at least 5 HLC hits among 7 adjacent DOMs on the same string within 1 μs. This trigger improves the sensitivity to vertical low-energy events. To also improve the sensitivity to horizontal low-energy events, specifically outside the DeepCore region, the 'volume trigger' was introduced. As a first step, this requires at least 4 HLC hits on non-DeepCore strings within 1 μs. For each of these hits a cylinder centered on this hit with a radius of 175 meters and a height of 75 meters (seven layers of DOMs) is defined. The trigger condition is satisfied when at least one of these volumes contains at least 4 hits.

Any overlapping trigger windows (either from the same or different trigger types) are combined. Then, all waveforms in a combined window plus the 4 μs before the first window and the 6 μs after the last window are saved together and called an 'event'. The event rate for all triggers combined is approximately 2500 Hz, and almost all of these events are due to atmospheric muons triggering the detectors. Due to this high rate, approximately 10% of the events are coincident events, where there are multiple visible particles in the detector at the same time which originate from different air showers. As will be shown in chapter 4, these events are complicated to reconstruct and therefore constitute an important background.

Though for some analyses all events are interesting, due to bandwidth constraints it is impossible to send them all to the Northern Hemisphere. For this reason, a further selection step is necessary, which consists of in-situ filters. In contrast to the triggers, which are only based on the geometrical and timing information of hit DOMs, most filters are based on initial reconstructions of the events. The input for these reconstructions is not the relatively complicated waveforms that are sent to the surface, but objects called 'pulses' which are described by a charge Q and a time t. These pulses are extracted from the digital waveforms after calibration, using an unfolding procedure with the waveform shape expected from one PE as a base [164].

The reconstructions which are applied will be discussed in more detail in section 4.1. Using their output, a number of filters dedicated to selecting specific event types (for example upgoing muon tracks, cascades or low-energy events) are applied. All events which pass at least one filter are then transferred North. The filters used for the analysis described in this thesis are discussed in section 4.4.1. In the North, base processing is applied to all events that are transferred. This processing consists of a number of further reconstructions of the events and the calculation of event variables.

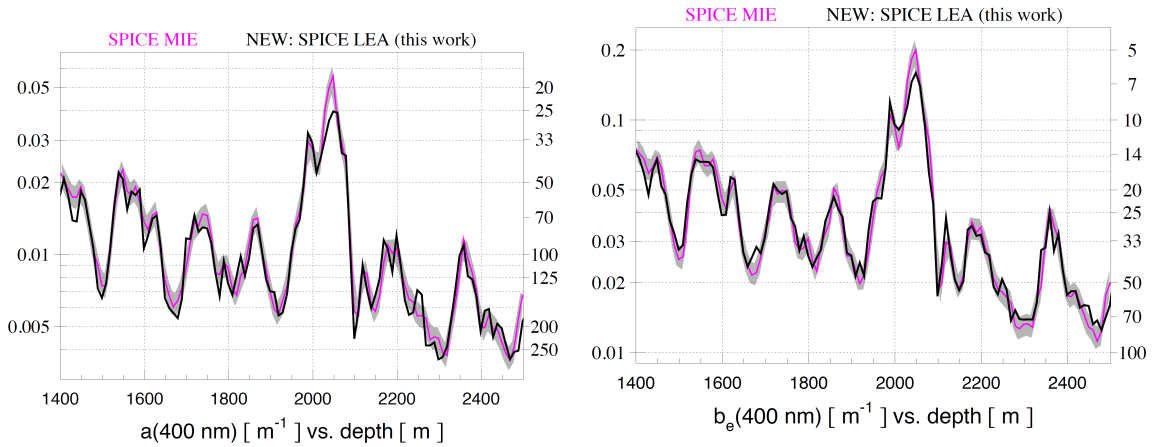
3.3.4. Optical properties of the South Pole ice

One important aspect of the IceCube detector which has not been discussed yet is the detection medium, which consists of the South Pole ice far below the surface. It originates from snow that fell more than 25,000 years ago. Each layer of snow added more weight on top of the layers below, compressing these until ice was formed. Currently, the ice cap at the South Pole is approximately 2.8 km thick. Any air bubbles or dust particles that were in the snow remained there while it was compressed, and could have a large influence on the

sensitivity of the detector. This is because the Cherenkov photons used for detecting and reconstructing particles can scatter off these bubbles and dust particles or be absorbed by them.

At the depths where the IceCube DOMs are located, the effect of air bubbles is negligible. This is due to the fact that the pressure has been high for so long that the air bubbles have become part of the crystal structure of the ice. As a consequence, they have an index of refraction almost equal to that of the ice, and the Cherenkov photons pass through them without significant scattering [181]. This means that the absorption and scattering lengths in the ice inside IceCube are determined by the concentration of dust, which varies throughout the detector, mainly as a function of depth. The variation is due to different ice layers being produced at different times in the history of the Earth, with different dust concentrations in the atmosphere.

To be able to correctly predict the behaviour of Cherenkov photons in IceCube and reconstruct particle properties from the photons detected by the DOMs, it is important to know the scattering and absorption lengths at each point in the detector as precisely as possible. To determine those, each of the DOMs contains a number of LEDs, and during special 'flasher runs' these are used to study the behaviour of known photons in the ice. Because the scattering and absorption lengths have the same order of magnitude, neither can be neglected, and there are no analytic formulae for the behaviour of the Cherenkov photons. Instead, the flasher runs are simulated using different 'ice models' (containing values for the scattering and absorption lengths at each depth), searching for the one which best fits the data [182]. Recently, it has become clear that the South Pole ice is anisotropic, meaning that its properties are different for different azimuthal directions. A new fit was done taking this anisotropy into account [183], and simulated events based on this fit are used for the analysis described in this thesis.



(a) The absorption length in meters (right axis) or its inverse (left axis) as a function of the depth.

(b) The effective scattering length in meters (right axis) or its inverse (left axis) as a function of the depth.

Figure 3.7.: Comparing the absorption and effective scattering lengths from the fit including the ice anisotropy (in black) to the fit where anisotropy is not taken into account (in pink, with the uncertainty band in grey). Taken from [183].

The resulting scattering and absorption lengths as a function of depth are shown in figure 3.7. Here the absorption length is the average length that a photon travels in the ice before it is absorbed. The scattering length shown here is the effective scattering length, which is defined as

$$l_e = \frac{l}{1 - \langle \cos\theta \rangle} \tag{3.7}$$

with l the geometrical scattering length (the average distance between two successive scatterings) and $\langle \cos\theta \rangle$ the average value of the cosine of the scattering angle. This effective scattering length is used because the distribution of scattering angles is not uniform.

Figure 3.7 shows that the scattering and absorption lengths vary significantly within the detector. The largest fluctuation is at a depth of approximately 2.1 km, where the scattering and absorption lengths are much smaller than at higher or lower depths. This is called the dust layer, since the dust concentration is higher there, and it is problematic for the reconstruction of events and vetoing of atmospheric muons. The dust layer is also shown in figure 3.6b, from which one can see that the main part of DeepCore is built just below this layer, where the ice is clearest.

Another feature, which is not part of the general ice model but determined separately, is the so-called hole ice: around each string, there is a column of ice that was refrozen after the hole for the string was drilled. The air bubbles in this hole ice have not yet become part of the crystal structure of the ice, so the scattering length is approximately 50 centimeters, much smaller than in the rest of the ice (the 'bulk ice'). However, in most cases the Cherenkov photons only travel through this hole ice for a short distance (unless they travel vertically).

It should be noted that scattering also has positive effects when studying downgoing events: as the PMTs are pointing downwards, the only way to detect significant amounts of light for downgoing tracks is if the photons are scattered.

3.4. Monte Carlo simulation

To be able to tell whether the IceCube data contains any neutrino events coming from dark matter annihilations, it is necessary to know what both these signal events and background events would look like. For this, Monte Carlo simulation is used, because the combination of initial fluxes, interactions, particle propagation in the ice and detector effects is too complicated to be described analytically. Though both atmospheric muons and atmospheric neutrinos (the main backgrounds discussed in section 3.2) are produced in the same process (cosmic ray showers in the atmosphere), they are simulated separately.

For the analysis described in this thesis, three types of Monte Carlo datasets are used: atmospheric muons, low-energy neutrinos (up to 195 GeV) and high-energy neutrinos (starting from 100 GeV). The neutrino datasets can be used to describe both the atmospheric neutrino background and the signal neutrinos from dark matter annihilations, by changing the per-event weight such that the weighted spectrum is correct. For the three types of simulation, the event generation differs, but the propagation of charged particles and photons through the ice and the detector simulation is the same.

As will be discussed in more detail in chapters 4 and 5, the atmospheric muon and atmospheric neutrino simulation is only used for illustrative purposes and validation. The event selection is determined using experimental data as the background sample, since the experimental data is highly dominated by atmospheric muons. For the final analysis, scrambled experimental data (events which have been given a random azimuth value) is used as background. In both cases, this is done to avoid possible complications or false signals due to the fact that Monte Carlo simulation is not perfect. It is still important to check that the signal simulation is correct, as errors would result in wrongly reconstructed dark matter properties if a signal is found, or wrong limits if it is not found. This will be done in chapter 4 by checking whether the atmospheric neutrino simulation agrees with the data, since the same datasets are used for the signal.

3.4.1. Atmospheric muons

For this analysis, the atmospheric muon background datasets are created by fully simulating cosmic ray interactions in the upper atmosphere and tracking only the muons. All other shower particles are ignored because they (except for the neutrinos, which are simulated separately) are incapable of reaching the IceCube detector, due to the 1.4 km of ice in between.

The cosmic ray showers are simulated using an IceCube-specific version of the COsmic Ray SImulations for KAscade package (CORSIKA) [184]. Events are generated following a power-law energy spectrum and then reweighted so they follow the composition and energy spectrum from the H3a model by Gaisser [170]. To be certain that all relevant energies are covered, cosmic rays with primary energies between 600 GeV and 10^{11} GeV are simulated. Below this energy range, the muons that are produced will not reach the IceCube detector, while above this energy range, the flux is so low that the probability of such an event happening during the livetime of IceCube is small. As muons from cosmic ray showers in the Northern Hemisphere will not reach IceCube, these showers are also not simulated.

The muons are propagated using a parametrization of their energy losses in the atmosphere and the upper layers of the ice, until they are close to the IceCube detector. From there onwards, they are propagated using the same software as is used for propagating the muons from neutrino interactions, which will be discussed in section 3.4.3.

3.4.2. Neutrinos

As neutrinos interact so rarely, it would be a waste of resources to simulate a flux and wait to see how many of those end up as visible events in or close to IceCube. Instead, neutrinos of a particular flavour are injected at the surface of the Earth following some energy and zenith distribution. They are then propagated through the Earth, taking into account possible interactions during propagation. Finally, each simulated neutrino is forced to interact in or close to the detector, and then the probability that such an interaction would happen is calculated and taken into account in the event weight.

High-energy neutrinos (from 100 GeV to 10^8 GeV) are simulated using a software package called neutrino-generator (often shortened to NuGen), which is based on the ANIS package [158]. It assumes that all interactions are DIS interactions (see section 3.1.2), which is a valid assumption at these energies, and uses parton distribution functions from CTEQ5 [185]. For low-energy neutrinos (from 3 GeV to 195 GeV), the GENIE event generator [186] is used,

which is valid down to lower energies than NuGen as it uses a model that also includes non-DIS interactions. In both cases, the Earth is described following the Preliminary Reference Earth Model [187].

Like the simulated atmospheric muon events, the neutrino events are initially generated randomly in azimuth and in the cosine of the zenith and follow a power-law energy spectrum. They are then re-weighted to either the atmospheric neutrino spectrum or the expected spectrum from dark matter annihilations. For the former, the spectrum from Honda et al. [171] is used, and then atmospheric neutrino oscillations (using the current world average values for the mixing parameters [11]) are applied.

Since the signal neutrinos are expected to come from a single point or small area on the sky, one might expect that the simulated signal dataset would consist of neutrinos which all have (almost) the same direction. However, as there are many possible point sources (not only for dark matter searches, but also for astrophysical neutrino searches), it would be too computationally expensive to produce such samples. Also, it is not necessary, because IceCube has a uniform azimuth acceptance due to the rotation of the Earth and because the effect of a zenith difference of a few degrees on the resulting event is small. Instead, as a signal sample one takes all events with a zenith less than 5 degrees away from the zenith of the target that is being considered.

The signal events are weighted to the flux from dark matter annihilations as given in formula 2.3. The J-factors given in table 2.1 are used, and dN_ν/dE (the expected energy spectrum for a single annihilation) is taken from simulations done with Pythia 8.1 [153] (some example spectra are shown in figure 2.15). After this, neutrino oscillations in vacuum are taken into account, as was discussed in section 2.5.1. The result of this is that neutrinos of all flavours will arrive on Earth, independent of the annihilation channel that is considered. As has been discussed in section 3.1.2, for this analysis track events are considered, which consist mainly of charged-current muon neutrino interactions. In about 17% of the charged-current tau neutrino interactions, the tau decays to a muon. However, since the muon energy will be significantly lower in this case, only muon neutrinos are considered for the simulated signal sample.

For both the atmospheric and the signal neutrino event samples, the low-energy (GENIE) and high-energy (NuGen) datasets are combined. This is done by multiplying the weight of each event in the overlap region (between 100 GeV and 195 GeV) in both samples with an extra factor. For GENIE, this factor is 1 at 100 GeV, 0 at 195 GeV and changes linearly with energy in between. For NuGen, it is the other way around: the factor is 0 at 100 GeV and 1 at 195 GeV. In this way, the final rates are correct and the change from one sample to the other is smooth.

3.4.3. Propagation

For the analysis described in this thesis, only muon neutrinos are considered. In this case, the products of the initial neutrino interaction are either a hadronic cascade and a muon (for CC interactions) or a hadronic cascade and an (invisible) neutrino (for NC interactions). The treatment of hadronic cascades differs for the Genie and Nugen datasets. For the former, hadrons below 30 GeV and electrons and photons below 100 MeV are directly propagated

using GEANT4 [188], while parametrizations are used for higher-energy particles. For the NuGen datasets, parametrizations are used at all energies. As photons usually scatter at least once before reaching a DOM, the exact development of the cascade is not visible, so parametrizations work well. However, at low energies the fluctuations between cascades get larger (as was discussed in section 3.1.2) and since there are less particles the computational cost of direct propagation is smaller, which is why direct propagation was introduced for the Genie datasets.

The muons coming out of the initial generators (either CORSIKA, NuGen or GENIE) are further propagated through the ice using the Muon Monte Carlo (MMC) package [189]. This simulates the muon energy losses discussed in section 3.1.3 using parametrizations. It also simulates the Cherenkov light that is produced during propagation (both from the original muon and from charged particles produced in interactions).

Once the Cherenkov emission from all charged particles is known, the photons need to be propagated through the ice, including scattering and absorption (using the ice model which was discussed in section 3.3.4). This is done by individually tracking each of the photons that are produced, to determine whether and when they arrive at a DOM, using the CLsim package. This is a rewrite of the Photon Propagation Code (PPC) package [190] and uses the same algorithms.

3.4.4. Detector simulation

Once it is known how many photons arrive at each DOM and when, the next step is to produce digitized waveforms which are comparable to those found in real data. This is done by first reproducing the behaviour of the DOM, for which one needs to take into account the angular and wavelength acceptances and the quantum efficiency of the PMT. These have been measured under laboratory conditions by illuminating the PMTs with different amounts of light, starting from single photons [178]. The detector is also regularly re-calibrated using the LEDs that are part of each DOM, which can produce single photon signals.

The next step is to include pulses produced by other sources, for example noise from electronics and afterpulses. The former are mostly due to thermal background and radioactive decays in the PMT, while the latter are produced by ions from residual gases in the PMT accelerating back to the photocathode and creating new electrons there [191]. After this, the digitization step (discussed in section 3.3.2) is reproduced, and from that point onwards simulation and data can be treated in the same way.

4. Event reconstruction and selection

As discussed in the previous section, only a small fraction of all events that trigger the IceCube detector are sent North. This small fraction is still highly dominated by background events, so the first step of this analysis is to apply an event selection to reduce this background. There are two types of background events: atmospheric muons and atmospheric neutrino. The event selection is focused on removing atmospheric muon events, as the atmospheric muon rate is much larger than the atmospheric neutrino rate and atmospheric neutrinos are hard to distinguish from signal neutrinos.

To remove atmospheric muons from the data sample, the two differences between atmospheric muons and (atmospheric or signal) neutrinos which were briefly discussed in section 3.2.1 are used. First, atmospheric muons are downgoing while the signal neutrinos considered in this analysis are upgoing. Second, atmospheric muons always enter the detector from outside while neutrinos can interact inside the detector. If all events were reconstructed perfectly, the first difference would be enough to get rid of all atmospheric muons, but this is not the case. The rate of atmospheric muons reconstructed as being upgoing is of the order of 100 Hz while the atmospheric muon neutrino rate is of the order of 10 mHz. For this reason, many methods have been developed to remove badly reconstructed events. These methods also remove badly reconstructed signal events, but these would not contribute much to the final analysis either way (as they do not appear to originate from the target).

Due to their importance to the event selection, this chapter starts with a thorough discussion of different track and energy reconstructions, including estimates of the error on the reconstructed direction, in section 4.1. Some of the event variables that are used in the event selection are derived directly from the track reconstructions. The other variables used in the event selection are described in section 4.2. Part of the event selection consists of cutting on one variable at a time, but the later levels consist of Boosted and Randomized Decision Tree Forests. These are algorithms combining the information encompassed in several event variables, and they are described in section 4.3. In section 4.4, the steps of the event selection used for this analysis are described. Finally, in section 4.5, the properties of the final event sample, which is the result of applying the event selection to the experimental data sample, are discussed.

The event selection described in section 4.4 (after the first step, which consists of a number of general filters) was developed specifically for this analysis and optimized on the signal expected from dark matter annihilations. In addition, some of the event variables (the ones in section 4.2.6) were developed specifically for this analysis, and some others were used in an event selection for the first time (for example the error estimate of the Boot reconstruction).

4.1. Event reconstruction

To perform the analysis described in chapter 5, there are three main event properties that need to be reconstructed. These are the energy and the direction of the muon, which are correlated to the energy and the direction of the parent neutrino, and the estimated error on the direction of the muon track. For all three there are several reconstruction algorithms available which are discussed here. As only charged current muon neutrino events are considered for this analysis, all algorithms described here assume that the event is a single muon passing through the detector. In most cases it is also assumed that both the starting and the stopping point of the muon are outside the detector, so the algorithms are reconstructing an infinite track. The ice model used in some of the reconstructions mentioned below has no anisotropy included (see section 3.3.4 for a discussion of anisotropy in the ice model).

As discussed in section 3.3.3, reconstruction algorithms are not run on the full waveforms as detected by the PMTs, but on a series of pulses (also often called 'hits') which are characterized by their charge Q (in units of PE, which stands for photo-electrons), time t and location $\vec{x}$ (the location of the DOM). For track reconstructions, the time and location are most relevant, while energy reconstructions include the charge as well. To remove pulses which are due to noise instead of Cherenkov photons, the pulse series is cleaned before it is used for the reconstructions.

4.1.1. Pulse series cleaning

Due to the PMT noise rate, which is of the order of 500 Hz, approximately 20 to 30 DOMs register a hit in a typical trigger window even though they did not detect any Cherenkov light. These can have a significant effect on the output of a track reconstruction, so it is important to remove them. For this, there is a standardized cleaning whose output is the input to most of the reconstructions that will be discussed in this section.

The first step of this standard cleaning is based on the RT radius. A pulse is within the RT radius of another pulse if they are within 1 μs and within 150 meters of each other. These values have been optimized to keep as many physics pulses (pulses caused by Cherenkov light) as possible while removing as many noise pulses as possible [192]. The starting sample for this algorithm consists of all HLC pulses that have at least two other HLC pulses within their RT radius. Then all (HLC or non-HLC) pulses within the RT radius of one of the pulses that is already selected are added to the sample, and this is repeated twice. The idea behind this method is to keep pulses that are correlated in time and space to multiple other pulses in the detector, which makes it less likely that they are due to noise.

For IC86-2012 and IC86-2013, the hit cleaning ended here and the resulting pulse series was used for the reconstructions. For IC86-2011, there was a second step, where it was determined which 6 μs window in the event contained the largest number of pulses, and all pulses outside this window were removed. This was done because it takes a relativistic muon less than 5 μs to cross the detector, so all pulses outside that window are likely to be due to noise. However, it turned out that in most cases these pulses are removed by the first step, so this second step had no effect. The only events where this second step had a significant effect were coincident events (where multiple independent particles are passing through the detector at the same time). However, in that case there is no guarantee that only the pulses caused by the most

interesting particle are kept by this rather crude method. More sophisticated methods to deal with coincident events are discussed in section 4.2.5.

4.1.2. Track reconstruction

Reconstructing the track means that both the direction (the azimuth and zenith) and the location (the x, y, z and time coordinates of some point along the track) are reconstructed. In most cases, the speed of the muon is set to its true value c (the speed of light in vacuum), but some track reconstructions also determine the speed of the muon.

In this section, several track reconstructions are discussed. Apart from some specialized reconstructions (Boot and FiniteReco), the main difference is the treatment of the ice properties. These have a large influence on the result of the reconstruction, and in general a better description of the ice results in a better track reconstruction. As these better reconstructions take more time to run, it is not feasible to run them on all events that pass one of the filters. Also, by running simpler reconstructions first, the result of these reconstructions can be used as a seed for the more complicated reconstructions. In general, the result of a reconstruction will be referred to by the name of that reconstruction (so 'LineFit' refers to the result of the LineFit reconstruction).

LineFit

The first and simplest track reconstruction that is run is called LineFit. It does not take into account the scattering and absorption in the ice or the shape of the Cherenkov cone. The basic assumption is that a muon produces a plane wave which travels parallel to the muon itself. In that case, it is possible to fit a track by minimizing the sum of the squares of the distances between the track and the hits. This sum is calculated as

$$\chi^2 = \sum_{i=0}^{n_{hits}} (\vec{x_i} - \vec{x_0} - \vec{v}t_i)^2, \tag{4.1}$$

where $\vec{x_i}$ and t_i are the location and time of the hit, $\vec{x_0}$ is the location of the muon at $t = 0$ and $\vec{v}$ is the velocity of the muon. Equation 4.1 can be minimized analytically, resulting in

$$\vec{x_0} = \langle \vec{x_i} \rangle - \vec{v}\langle t_i \rangle \text{ and } \vec{v} = \frac{\langle \vec{x_i} t_i \rangle - \langle \vec{x_i} \rangle \langle t_i \rangle}{\langle t_i^2 \rangle - \langle t_i \rangle^2}. \tag{4.2}$$

The speed of the muon is not fixed. As the muons passing through the IceCube detector generally have a speed approximately equal to c, the magnitude of $\vec{v}$ can be used to remove badly reconstructed events (for which $\vec{v}$ can be very different from c).

With this method, a few hits which do not agree with the assumption that all hits are close to the true track can have a large influence on the result. Even after the hit cleaning described in section 4.1.1, many events still contain a few noise hits or hits from highly scattered Cherenkov photons. For this reason, equation 4.1 is altered such that each hit has a weight, and hits that are qualified as outliers are given a lower weight. The definition of which hits are 'outlier hits' is based on studies using Monte Carlo simulation [193].

Maximum likelihood fits

The LineFit algorithm described above can be improved by taking into account the Cherenkov light profile and scattering and absorption in the ice. Since scattering and absorption are probabilistic processes, the problem can no longer be solved analytically. Instead, a maximum likelihood method is used, which is discussed in detail in [194]. The basic idea is to define the likelihood $\mathcal{L}$ to get a set of observed data points $\{\vec{r_i}\}$ (in this case, the locations and times of the hits) given a set of parameters $\vec{\theta}$ (in this case, the track parameters). For independent data points, this likelihood is given by

$$\mathcal{L}(\{\vec{r_i}\}|\vec{\theta}) = \prod_i p(\vec{r_i}|\vec{\theta}) \tag{4.3}$$

with $p(\vec{r_i}|\vec{\theta})$ the Probability Density Function (PDF), which is the probability to get a data point $\vec{r_i}$ given a set of (track) parameters $\vec{\theta}$. One can then maximize $\mathcal{L}$ (or, as is usually done, minimize the negative logarithm of $\mathcal{L}$) to get the best fit for the parameters $\vec{\theta}$. Usually, the parameter space for $\vec{\theta}$ is too large to test all possible values, so a seed is used as a starting point and a good minimizing algorithm is necessary. For the reconstructions described in this chapter, the MINUIT algorithm is used [195].

For the track reconstruction problem considered here, $p(\vec{r_i}|\vec{\theta})$ is the probability to detect a photon at a certain location and time if it was produced by a muon following the track described by $\vec{\theta}$. When considering a DOM with a fixed position, the probability $p(t_{res}|\vec{\theta})$ for a photon to have a certain time residual t_{res} can be used instead. The time residual is the difference between the arrival time of the photon at a DOM and the time at which the photon would have arrived if it had travelled in a straight line from the track to the DOM, without being scattered.

Equation 4.3 describes the likelihood when all hits are taken into account. A possible improvement is to only consider the first hit in each DOM, which is usually the least scattered, so for track reconstructions it is the most important. This is called the Single Photo Electron (SPE) likelihood, which is defined as [194]

$$\mathcal{L}_{SPE}(\{\vec{r_i}\}|\vec{\theta}) = \prod_{i=1}^{n_{firsthits}} p(t_{res,i}|\vec{\theta}). \tag{4.4}$$

In the SPE likelihood, all first hits are treated equally. Another possible improvement is to take into account how many hits a DOM has (how many photons were detected in that DOM). This is useful because the first photon in a DOM that detected multiple photons is on average less scattered than the first photon in a DOM which detected no other photons. For this the Multiple Photo Electron (MPE) likelihood is used, which is defined as [194]

$$\mathcal{L}_{MPE}(\{\vec{r_i}\}|\vec{\theta}) = \prod_{i=1}^{n_{firsthits}} n_i \cdot p(t_{res,i}|\vec{\theta}) \left(\int_{t_{res}}^{\infty} p(t|\vec{\theta})\right)^{n_i} \tag{4.5}$$

with n_i the number of hits in the DOM. The MPE and SPE likelihoods are equal for $n_i = 1$, as can be seen by comparing equations 4.4 and 4.5.

The time residual PDF is mainly determined by scattering in the ice. However, other effects like PMT noise and PMT jitter (uncertainties on the arrival time as measured by the PMT) cannot be ignored since unlike scattering, those can lead to negative t_{res} values. The noise is taken into account by adding a flat noise probability to the PDF, for which a rate of 10 Hz is assumed. The PMT noise rate is much higher (as discussed in section 3.3.2), but most of these noise hits have already been removed by the hit cleaning procedure discussed in section 4.1.1. To take into account the PMT jitter, the likelihood is convolved with a Gaussian with a width of 4 ns [196].

For both the SPE and the MPE likelihood, the likelihood value of the best-fit track is a measure for how well this track fits the data. To avoid dealing with small numbers, usually the negative logarithm of the likelihood value is used, which is referred to as 'logl'. Its value depends on the number of hits that are considered, as the likelihood is the product of probabilities for each of these hits. To be able to compare the reconstruction quality of different events (with different numbers of hits), one could divide the logl by the number of degrees of freedom. This is the number of data points (for a track fit, this is the number of hit DOMs) minus the number of parameters that is being fit (for a track fit, this is five). However, in earlier studies done within the IceCube collaboration it was found that under some circumstances a slightly different value performs better in separating signal and background, so we define a general reconstruction quality parameter as

$$\text{rlogl}n = \frac{-\log(\mathcal{L})}{N_{ch} - n}, \tag{4.6}$$

where N_{ch} is the number of hit DOMs, n could have different values (e.g. five or three) and rlogl stands for 'reduced log-likelihood'.

The Pandel distribution

To be able to perform the track fit described in the previous section, one needs to know the time residual PDF for all possible orientations, depths and distances between the track and the DOM. Using laser light to study light propagation in the water of lake Baikal, it was found that the time residual PDF can be approximated by the so-called Pandel distribution, which is defined as [197]

$$p_P(t_{res}|\vec{\theta}) = \frac{1}{N(d)} \frac{\tau^{-(d/\lambda_s)} \cdot t_{res}^{d/\lambda_s - 1}}{\Gamma(d/\lambda_s)} \cdot \exp\left(-t_{res} \cdot \left(\frac{1}{\tau} + c_{medium}\lambda_a\right) - \frac{d}{\lambda_a}\right), \tag{4.7}$$

with

$$N(d) = e^{-d/\lambda_a} \cdot \left(1 + \frac{\tau \cdot c_{medium}}{\lambda_a}\right)^{-d/\lambda_s}, \tag{4.8}$$

with d the distance between the track and the DOM, λ_s the effective scattering length, λ_a the absorption length, Γ the gamma function, c_{medium} the speed of light in the given medium and τ a parameter which needs to be fitted. In IceCube, to take into account the angular dependency of the DOMs (since the PMTs point downwards), the distance d is replaced by an effective distance which is a function of the zenith of the track [194].

As has been discussed in section 3.3.4, the effective scattering length and absorption length depend on the depth. However, from comparisons to Monte Carlo simulations it has been

shown that for IceCube, the Pandel distribution performs well at all depths if the following parameter values are used [194]:

$$\lambda_s = 33.3 \text{ m}, \tag{4.9}$$
$$\lambda_a = 98 \text{ m}, \tag{4.10}$$
$$\tau = 557 \text{ ns}. \tag{4.11}$$

Though the Pandel distribution is a simplification of the true behaviour of photons in ice, it has the advantages that it is normalized and its integral can be calculated analytically. Examples of the Pandel distribution at different effective distances are shown in figure 4.1. Here it is clear (as can also be seen from equation 4.7) that for $d_{eff} > \lambda_s$, the shape of the distribution changes: the probability to detect unscattered photons becomes negligible.

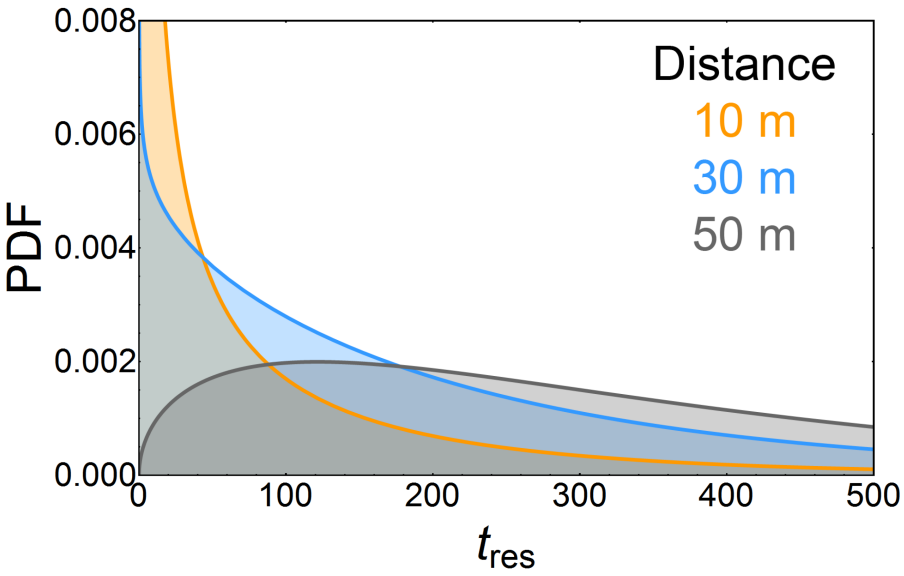


Figure 4.1.: The Pandel distribution for three different effective distances (corrected for the angular dependency of the DOMs) between the track and the DOM. Taken from [198].

The Pandel distribution is used to construct both the SPE likelihood and the MPE likelihood. The resulting track reconstructions are usually referred to as PandelSPE and PandelMPE. In the IceCube base processing, PandelSPE is seeded with LineFit and PandelMPE is seeded with PandelSPE.

Using spline fits to tables

The PandelSPE and PandelMPE fits perform well for most events, as will be shown later on, but there is room for improvement. Instead of using a simplified analytical function to describe the time residuals at all depths, it would be better to correctly describe the behaviour of the Cherenkov photons and take into account the different absorption and scattering lengths at different depths.

Since detailed Monte Carlo simulations using the current best knowledge of the ice are available, it makes sense to extract the time residual PDF from these. The results of these simulations have been tabulated and these tables were fitted using a linear combination of basis splines [199]. Just like the Pandel distribution, these splines can be convolved with a Gaussian to take into account the PMT jitter. The resulting reconstructions are called SplineMPE and SplineSPE, and are not yet part of the IceCube base processing. Since they are time-consuming, they are only performed after a number of preliminary cuts have been applied to reduce the data rate. For the analysis described in this thesis, SplineMPE was used, seeded with PandelMPE and BootMPE (which is discussed in the next section).

Bootstrapping

The reconstructions described above together form the reconstruction chain used in many IceCube analyses, where each reconstruction is seeded with the result of the previous one. However, one track reconstruction (referred to as 'Boot' in this thesis) follows a different approach. The first difference is the seed: Boot creates its own seed by forcing the track to go through the center of gravity of the event (the charge-weighted average position of all hits) and then testing a large number of possible directions to see which gives the best likelihood value.

The main difference with the standard reconstructions is that Boot uses bootstrapping to get a more accurate reconstruction. The general idea of bootstrapping is to use random sampling with replacement to get better estimates of the properties of a distribution, and the error on these estimates [200]. In the Boot reconstruction, this is applied in the following way: first, the event is reconstructed (using the Pandel distribution, the SPE or MPE likelihood and the seed described above) using the original pulse series. Then, four new pulse series are created from the original pulse series using random sampling with replacement, and each of these is also reconstructed. Finally, the event is reconstructed using the original pulse series again, but now the seed is the average of the reconstructions of the four new pulse series. The final result is the reconstruction with the highest likelihood value out of all reconstructions that were done. This result is called BootMPE or BootSPE (depending on whether the SPE or MPE likelihood was used).

Boot was originally designed to give an error estimate on the direction reconstruction, σ_{Boot}, which is determined from the variation of the reconstructed directions using the new pulse series. This error estimate performs significantly worse than the error estimate discussed in section 4.1.3, but as will be shown later on, it is useful for removing badly reconstructed events from the data sample.

Finite tracks

The track reconstructions discussed so far all assume that the muon track is infinite. However, for the energy range considered in this analysis this is often not the case, as for muon energies below a few hundred GeV the muon track length is smaller than the size of the detector (see figure 3.2). For the direction reconstruction this is mostly irrelevant, but there are two reasons why it is useful to reconstruct the (possible) starting and stopping points of the muon. First, one of the main differences between atmospheric muons and muons produced in neutrino interactions is that the former always enter the detector from outside while the latter can also be produced inside the detector (resulting in so-called starting tracks). Second,

the range of a low-energy muon is related to its energy by equation 3.6, so if the range can be determined, it can be used in the energy reconstruction (as is discussed in section 4.1.4).

Reconstructing the starting and stopping points together with the other track parameters (position and direction) is complicated and time-consuming, due to the large parameter space. For this reason, the algorithm that is used (FiniteReco, which is discussed in detail in [201]) assumes the track as reconstructed by PandelSPE is correct, and only reconstructs the starting and stopping points along that track.

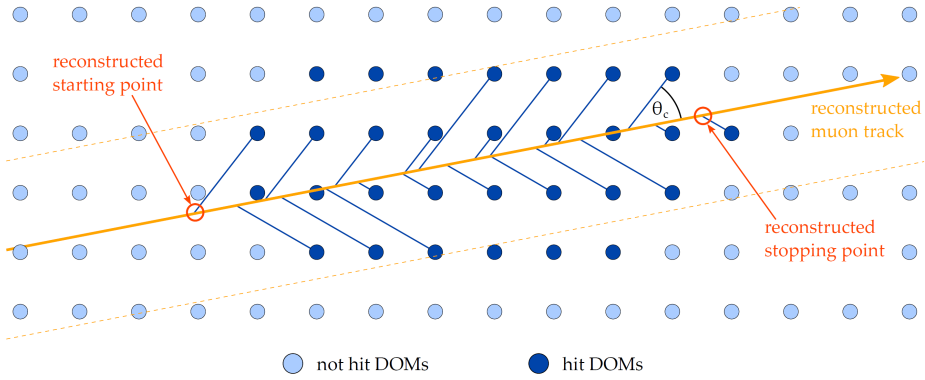


Figure 4.2.: The construction of the first-guess starting and stopping points along a muon track. Taken from [201].

FiniteReco consists of two steps: first all DOMs which are less than 200 meters from the PandelSPE track are projected onto the track using the Cherenkov angle. The first guess for the starting and stopping points are the outermost projections, as is shown in figure 4.2. This first guess is then used as a seed for a likelihood fit, for which only the DOMs projected before the first-guess starting point or after the first-guess stopping point are used. Due to the way the first guess was constructed, none of these DOMs detected a hit, and using Monte Carlo simulation one can calculate what the probability for this would be for different starting and stopping points. Maximizing the likelihood will then give the final fit results for the starting and stopping points, from which the muon track length can be calculated. The typical accuracy of the starting and stopping points is approximately 25 meters [202], which is less than the distance between the IceCube strings.

Performance

The performance of the reconstructions discussed above (except for FiniteReco, which does not reconstruct the direction) is compared in figure 4.3, using a muon neutrino simulation dataset. The median angular resolution (the median angle between the reconstructed direction and the true neutrino direction) is plotted as a function of neutrino energy. For this, the event sample at the penultimate level of the event selection discussed below is used, since this is the level where the more time-consuming reconstructions (SplineMPE, BootMPE and BootSPE) are applied. The median resolutions are also compared to the median kinematic angle, which is the angle between the true neutrino direction and the true muon direction. If it were possible to reconstruct the direction of the muon perfectly, the median angular resolution would be equal to the median kinematic angle.

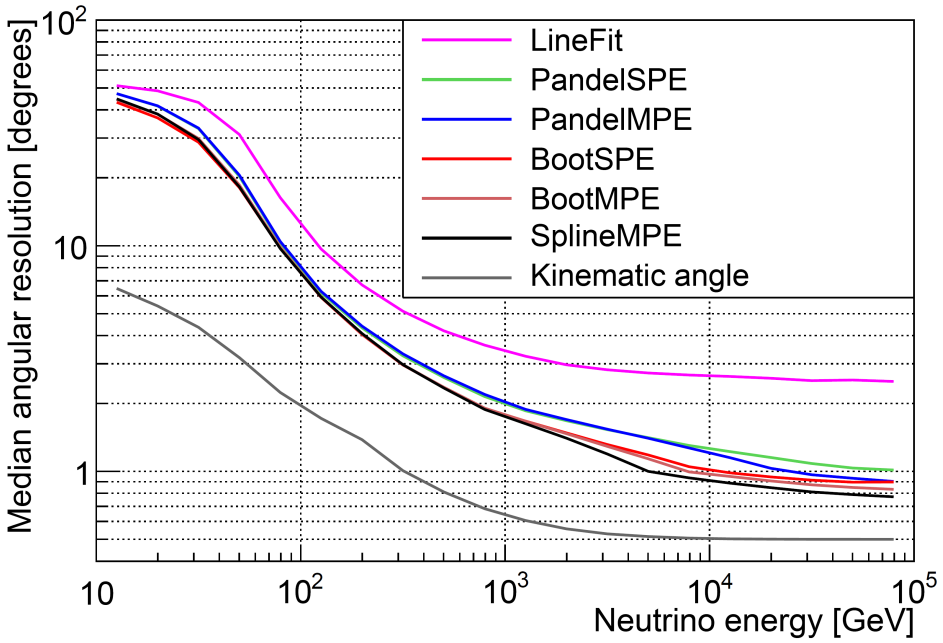


Figure 4.3.: Comparison of the median angular resolution as a function of neutrino energy for the track reconstructions discussed in section 4.1.2. A muon neutrino simulation dataset at the penultimate level of the event selection is used.

As expected, LineFit performs worst, followed by PandelSPE and PandelMPE, and finally BootMPE, BootSPE and SplineMPE. This is because the descriptions of the ice get increasingly better, and each reconstruction uses the result of the previous as a seed. Boot is somewhat of an outlier since it uses the Pandel distribution, but has resolutions comparable to SplineMPE. This is due to the different method employed by Boot, and it would be interesting to see whether this bootstrapping could be combined with the Spline likelihood.

One can see that for low-energy events (below a few TeV) the difference between MPE and SPE is negligible. This is explained by the fact that the MPE likelihood reduces to the SPE likelihood if each DOM detects only one hit, which is the case for most low-energy events. For high-energy events on the other hand, MPE performs better than SPE, since it takes into account more relevant information.

For the analysis described in this thesis, SplineMPE was used as track reconstruction for the final analysis. Though BootMPE and BootSPE perform slightly better at low energies, the estimated error (which is the topic of the next section) for SplineMPE performs significantly better than the estimated error for BootMPE and BootSPE. Since both the direction reconstruction and its estimated error play an important role in the final analysis, the combination of SplineMPE and Paraboloid gives the best sensitivity. However, variables based on all reconstructions described above are used for the event selection, as will be shown in section 4.4.

4.1.3. Estimated error of the direction reconstruction

There are two reasons why one would be interested in the estimated error of the direction reconstruction. The first is to use this to remove background events: if the estimated error

is large, that means that this is a badly-reconstructed event, so it is likely to be a down-going atmospheric muon. The second is to use this in the final analysis, where the goal is to determine how many events are coming from the direction of the target. If a per-event estimated error is available, one can give a larger weight to events which have a reconstructed direction close to the target and a small estimated error, than to events which also have a reconstructed direction close to the target but a large estimated error. The details of how this is taken into account can be found in section 5.1.

To determine the estimated error, the Paraboloid algorithm was used, which is described in more detail in [203]. This algorithm first calculates the negative log-likelihood (logl) values on a grid of zenith and azimuth values close to the result of the original fit, optimizing with respect to the other variables (the position of the track) for each grid point. In this way, a logl landscape is defined, which is expected to be rather flat around the best-fit point for badly-reconstructed events while it will have a sharp, well-defined minimum for well-reconstructed events. To quantify this, the landscape is fitted with a two-dimensional paraboloid, which results in a confidence ellipse (in most cases, zenith will be reconstructed better than azimuth). The overall error estimate, usually called Paraboloid sigma, is calculated as

$$\sigma_P = \sqrt{\frac{\epsilon_1^2 + \epsilon_2^2}{2}} \tag{4.12}$$

where ϵ_1 and ϵ_2 are the major and minor axes of the ellipse. Paraboloid can be used both with the Pandel and Spline time residual distributions, and also with both the SPE and MPE likelihoods. For this analysis, since SplineMPE is the track reconstruction used in the final analysis, Paraboloid uses the SplineMPE likelihood and best fit result.

In figure 4.4, a two-dimensional histogram of the reconstruction error (the angle between the reconstructed direction and the true direction) and the Paraboloid sigma is shown. The two values are clearly correlated, however it is also clear that on average, Paraboloid underestimates the true error. In section 5.2, a method for recalibrating the Paraboloid sigma is discussed, which solves this problem.

4.1.4. Energy reconstruction

For the analysis described here, the reconstructed energy is not used for the event selection. This means that energy reconstructions are only run on events which pass the full event selection, and it is less important how fast they are. Since low- and high-energy events are quite different, they are reconstructed using different methods.

High-energy events

The higher the energy of a muon is, the longer its track will be (see figure 3.2), so the higher the probability is that its starting point is not located inside the IceCube detector. If this is the case, it is impossible to reconstruct the energy of the original neutrino, as the hadronic cascade cannot be detected and the muon has already lost part of its energy before entering the detector. However, one can reconstruct the energy that the muon has when it passes through the detector, which is correlated to the energy of the original neutrino.

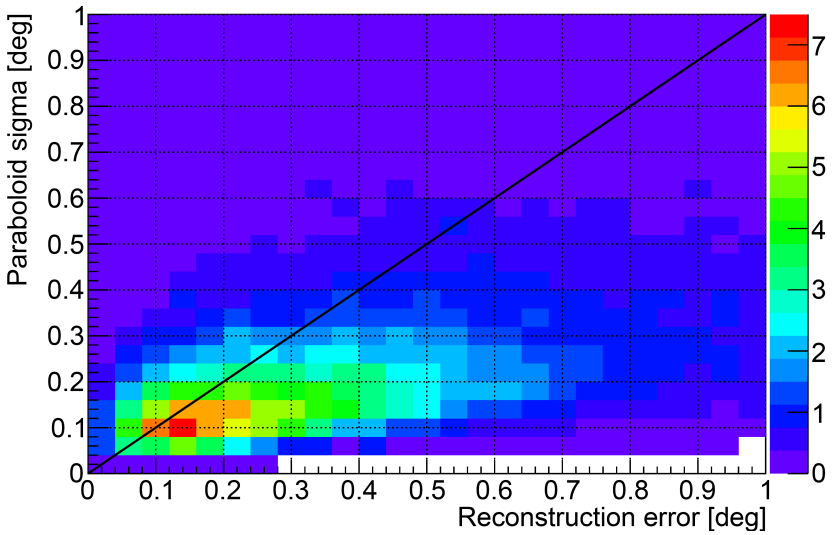


Figure 4.4.: Two-dimensional histogram of the reconstruction error and the Paraboloid sigma, using arbitrary units on the colour scale. A muon neutrino simulation dataset at the penultimate level of the event selection is used, and each event is weighted with E_ν^{-2}. The black line shows where both values are equal.

As has been shown in section 3.1.3, specifically equation 3.5, for muon energies above approximately 700 GeV the muon energy is proportional to its energy loss per unit distance dE_μ/dx. This is proportional to the number of Cherenkov photons that are radiated off [204]. To reconstruct the muon energy loss, an algorithm called MuEx is used. This is based on an analytical approximation of the behaviour of Cherenkov photons in ice [180], which is described below. As input, MuEx needs a reconstructed track, for which the SplineMPE track is used.

Up to distances of the order of 10 meters, where absorption does not play a role and there is a cylindrical symmetry, the photon density is proportional to $1/r$, where r is the distance to the track. At distances of the order of 100 meters and larger, MuEx uses the diffusive approximation where the photon paths can be approximated by a random walk (due to scattering). In that case,

$$p_\gamma \propto \frac{e^{\frac{-r}{\sqrt{\lambda_a \lambda_s/3}}}}{\sqrt{r}} \tag{4.13}$$

with p_γ the photon density, λ_a the absorption length and λ_s the effective scattering length. Since these lengths depend on the depth, the average values along the line between the track and the DOM are used. The two regimes are combined using an empirical function which matches the functions given above for small and large r. The resulting analytic formula for the photon density at a certain distance from the track is used in a likelihood method. This determines the muon energy loss dE_μ/dx which best fits the number of photons (the total charge of the hits) detected in the region around the track, taking into account that some hits are produced by noise.

The output of the MuEx algorithm is not a reconstructed neutrino energy, but a value which is correlated to the neutrino energy (an energy proxy). As the analysis described in

this thesis does not use the neutrino energy itself, but the difference in distributions between background events and signal events, this is not a problem.

Low-energy events

For events with neutrino energies below 700 GeV, the muon energy loss is approximately constant with energy (see equation 3.5). This means that the initial muon energy can be calculated from the range of the muon (as shown in equation 3.6). Also, for events with muon energies below a few hundred GeV, the initial neutrino interaction often takes place inside the detector (since the range is small), which means that the hadronic cascade can be detected as well. As has been discussed in section 3.1.2, the light output of a hadronic cascade is approximately linearly related to its energy, though there are large fluctuations. This means that the neutrino energy for low-energy events (neutrino energies below a few hundred GeV) can be reconstructed by determining both the energy of the muon (which can be calculated from its range) and the energy of the hadronic cascade. The sum of these is the reconstructed neutrino energy. This is the basic idea of the LEERA algorithm, which is described in more detail in [202].

As a first guess, the LEERA algorithm uses the starting and stopping points determined by the FiniteReco algorithm (discussed in section 4.1.2) as the neutrino interaction vertex and the muon stopping point. The neutrino interaction vertex is used as a seed for the second part of the method, while the stopping point is kept fixed. The goal of the second part is to reconstruct both the energy of the hadronic cascade and the position of the neutrino interaction vertex (which is the starting point of both the muon and the cascade) simultaneously. This is done by maximizing a likelihood, using the expected light output from both low-energy muons and cascades as determined from Monte Carlo simulations. It is assumed that the muon and the cascade both have the same direction and that the interaction vertex is located somewhere along the muon track.

Due to the assumptions made, LEERA only performs well for events where the initial neutrino interaction takes place inside the detector, with muon energies lower than a few hundred GeV. As has been discussed above, the MuEx reconstruction only performs well for neutrino energies above approximately 1 TeV. This means that there is a gap in between, where the muon energy loss is not proportional to the muon energy, and the neutrino interactions do not take place inside the detector. The energy of these events is hard to reconstruct, and no suitable algorithm has been developed yet, so usually one of the other algorithms is also used for this energy region.

4.2. Event variables

The goal of the event selection described in this thesis is to remove atmospheric muon events. To be able to reach this goal, event variables are needed which have different distributions for signal events and atmospheric muon events (background events). All event variables used in this event selection are based on one or multiple of the following principles:

- Atmospheric muon events are always downgoing while signal events are always upgoing. An example of an event variable based on this principle is the reconstructed direction, but there are also others which are directly calculated from the hit series.

- If an atmospheric muon event is reconstructed as upgoing (which means that it would survive cuts based on the first principle), that means that it is wrongly reconstructed. Many methods have been developed to tag misreconstructed events or quantify how well-reconstructed an event is.

- A variation on the previous principle is that it is also possible to tag events that are difficult to reconstruct, without looking at the quality of the reconstruction itself. An example of these are coincident events, where there are multiple particles in the detector at the same time, which means that the assumptions made by track reconstructions are wrong. Since these types of events have a high probability of being wrongly reconstructed, they are often removed from the sample.

- Atmospheric muons always enter the detector from outside, which means that they are already radiating off Cherenkov photons before they enter the detector. On the other hand, neutrinos can create a muon which starts inside the detector (called a 'starting track'), in which case the light from this muon would not hit the DOMs on the outer layer of the detector. This means that especially for low-energy events (with short muon tracks), the outer layers of the IceCube detector can be used as a veto: if light is detected there, this is most likely an atmospheric muon event.

In the remainder of this section, the variables used in the event selection described in section 4.4 are described. Unless noted otherwise, all of these use the hit series resulting from the cleaning discussed in section 4.1.1.

4.2.1. Event direction variables

The zeniths as reconstructed by several of the reconstructions discussed in section 4.1.2 are used in the event selection. All of these reconstructions use the hit series after the cleaning discussed in section 4.1.1. This hit series still contains some noise, so a stricter hit cleaning is also used, where two pulses are in each other's RT radius if they are within 75 meters and 625 ns (instead of 150 meters and 1 μs). For the analysis described in this thesis, the SPE likelihood and bootstrapping method are used in combination with this stricter hit cleaning, and the resulting reconstruction is referred to as StrictBootSPE.

In an upgoing event, the hits that are later in time are expected to have higher z values (which means that they are closer to the surface of the ice). There are several variables which test this assumption, for example

$$z_{trav} = \sum_{i}^{N_{hits}} \frac{z_i - \langle z_{1stquart} \rangle}{N_{ch}} \tag{4.14}$$

where the sum is over all hits and $\langle z_{1stquart} \rangle$ is the average z value of the first quartile of hits (when ordered in time). In other words, z_{trav} is the average difference (in meters) between $\langle z_{1stquart} \rangle$ and the hits in the hit series. For upgoing events, this will usually have a positive value, since $\langle z_{1stquart} \rangle$ will be small, while for downgoing events it will be negative.

A second variable is the so-called z pattern. This is calculated by ordering all hits in time and comparing the z value of each hit with that of the previous one. The z_{patt} value is the number of times that a hit has a higher z than the previous hit minus the number of times

that a hit has a lower z than the previous hit. Like for z_{trav}, in general a higher z_{patt} value means that the event is more likely to be upgoing.

4.2.2. First HLC hit

As has been discussed before, HLC hits have a lower likelihood of being due to noise than non-HLC hits. For this reason, the location of the first HLC hit can be used to remove atmospheric muon events, for which the probability is high that the first HLC hit is located either at a high z value or on the outer string layer. The variables that are being used here are z_{HLC}, which is the z value of the first HLC hit, and l_{HLC}, which is the string layer of the first HLC hit. These layers are defined as follows: layer 1 is the outermost layer of the detector (also known as the boundary), layer 2 is the next-to-outer layer, layer 3 is the next-to-next-to-outer-layer and layer 4 contains all other strings (including all DeepCore strings). For these variables, all HLC hits are used, including the ones that are removed by the hit cleaning described in section 4.1.1.

4.2.3. Direct hits

A direct hit is a hit produced by a photon which either does not scatter before it hits the DOM or scatters only slightly. In the analysis described in this thesis, hits are called direct if -15 ns $< t_{res} <$ 150 ns, where t_{res} is the time residual. This is the difference between the actual arrival time of the hit and the time at which it would have arrived if the photon had travelled directly from the reconstructed track to the DOM without scattering.

The time residual depends on the reconstructed track, as the photon travel times are calculated with respect to the track. If a track is badly reconstructed, even photons which do not scatter will have large (positive or negative) time residuals, because the time residuals are calculated with respect to the wrong track. This means that on average, the number of DOMs with direct hits is larger for well-reconstructed events. This value is referred to as N_{dir}. However, in some cases there are many hits close together. In that case, a track can pass through that area and N_{dir} can be large, even though the direction is reconstructed wrongly. For this reason, the direct length L_{dir} is also often used. This is calculated by projecting all direct hits on the track and taking the distance between the two that are furthest apart.

4.2.4. Track quality parameters

In addition to the number of direct hits and the direct length discussed above, there are several other reconstruction quality parameters. First, since there are several available track reconstructions (as discussed in section 4.1.2), a useful strategy is to compare the results of these track reconstructions. In general, if an event is easy to reconstruct, different track reconstructions will give similar results. On the other hand, if an event is difficult to reconstruct, the space angle between the two reconstructions (denoted as $\Delta\Psi(\text{reco}_1, \text{reco}_2)$) will in general be larger, so events with a large $\Delta\Psi$ are more likely to be misreconstructed. This holds in particular for reconstructions that are based on different methods, for example LineFit and SplineMPE.

As was already discussed in the subsection about direct hits, if the track is reconstructed correctly, most of the hits will be close to the reconstructed track, especially the high-charge hits. A way to measure this is to determine the charge-weighted average distance between

the hits and the track $\langle d_{ht} \rangle$.

A third track quality parameter is calculated by considering all DOMs in the cleaned hit series which are located closer than 150 meters to the track. L_{emp} is the maximum length along the track where no such DOMs are found. For well-reconstructed infinite tracks (meaning that the muon does not start or stop inside the detector) L_{emp} is usually small, as Cherenkov radiation is produced everywhere along the track. For badly reconstructed events, where the track only passes through part of the actual event, L_{emp} will on average be larger. This variable works better for high-energy signal events than for low-energy signal events, as well-reconstructed events with a short muon track length will usually have a large L_{emp}.

4.2.5. Event splitting

As was discussed in the introduction to this section, coincident events are hard to reconstruct, and if they can be found they can be removed from the sample. One possible method is to divide the hit series into two parts, reconstruct both of these separately, and determine the lowest zenith (the most downgoing) of the two reconstructions. Two types of division are used: the first one is called time splitting, here the average time of the hit series (using only the first hit in each DOM) is determined and used to split the hit series in an 'early' and 'late' subseries. Second, there is geometrical splitting, where the center of gravity of the hit series is determined and the hit series is split by the plane which passes through the center of gravity and is perpendicular to the track as reconstructed by PandelMPE. The goal of these splitting methods is to see whether there are subevents which are reconstructed as downgoing while the full event is not, which points towards coincident events. These coincident events could either consist of two atmospheric muons or of an atmospheric muon and a neutrino, but in both cases they are most likely badly reconstructed and need to be removed.

There is also a more sophisticated method of event splitting, which is called Topological Splitting. Because muons pass through the detector with speed c, the time difference between hits that are caused by the same muon should be approximately $\Delta t = R/c$, with R the distance between the two hits. For this algorithm, two hits are considered to be causally connected if $|\Delta t - R/c| < 1000$ ns. Additional criteria are that the horizontal distance between the hits should not be more than 300 meters and if the hits are on the same string, the vertical distance should not be more than 30 meters. The algorithm searches for causally connected hits (considering only the first hit in each DOM), and all hits which are causally connected to eachother form a subevent. As the cleaned hit series could still contain some noise, which often shows up as very small subevents, the number of subevents which contain more than 4 pulses within 4000 ns is considered, and referred to as N_{sub}. If this is larger than 1, it is likely that this is a coincident event.

4.2.6. Hit series properties

There are also other properties of the hit series which can make it hard to reconstruct the event, which makes it more likely that the event is badly reconstructed. Two of these are used here: first, one can determine Q_b/Q_{tot}, which is the fraction of the total charge of the event which is located on the outermost string layer of the detector (the boundary). Events where this fraction is high are usually events where the charged particle did not pass through the detector. As these events are hard to reconstruct correctly, it is likely that they are misreconstructed downgoing events instead of correctly reconstructed upgoing events. The

second variable is the fraction of the total charge in the event which is located on the string that has the highest charge (called Q_s/Q_{tot}). Again, these events are harder to reconstruct than events where the charge is more evenly divided over several strings.

4.3. Boosted and Randomized Decision Tree Forests

Boosted and Randomized Decision Tree Forests (usually referred to as BDTs) are algorithms combining the information encompassed in several event variables to separate signal and background events in large datasets. For the later levels of the event selection described in section 4.4, they perform significantly better than tuning cuts on single variables by hand (as is done for the earlier levels of the event selection). The cost is that the event selection is less easy to understand (the BDT is a 'black box') and that one should watch out for the problem of overtraining, which is discussed at the end of this section.

To train and test the BDT, event samples are needed which contain events that are known to be signal or background. For this analysis, simulated signal datasets (as discussed in section 3.4) are used as signal and a small fraction of the experimental data is used as background (since this is highly dominated by atmospheric muon events). More details will be given in section 4.4. Half of each sample is used for training the BDT (teaching the algorithm how to separate signal and background) and the other half is used to test the performance of the BDT after it has been trained. As input for the training, one also needs a number of event variables (for example the ones that have been discussed in section 4.2) which have different distributions for signal and background. Examples of distributions of variables that are used for BDTs at various stages of this event selection can be found in appendices A, B and C.

4.3.1. Decision trees

The basic building block of a BDT is a decision tree, an example of which is shown in figure 4.5. It consists of split nodes (the rectangles) and leaf nodes (the ellipses). At each split node, a cut on one event variable is used to separate the sample in two subsamples. Depending on whether the part of the training sample that ends up in a leaf node consists mainly of signal events or of background events, it is classified as a 'signal' or a 'background' leaf. Using this tree, one can classify an event of unknown type as 'signal' or 'background' by passing it through the tree to see whether it ends up in a signal or background leaf.

A decision tree performs well if the signal leaves have a high signal purity p (which is the ratio of the signal weight and the total weight) and the background leaves have a low p. For the analysis described in this thesis, the performance of a tree is measured by the Gini impurity [205], which for a sample with signal purity p is defined as $S_G(p) = p \cdot (1 - p)$. To determine the level of separation that results from applying the full tree,

$$S = \sum_i W_i S_G(p_i) \tag{4.15}$$

is used, where the sum runs over all leaf nodes, W_i is the total weight of all training events (signal and background) in leaf i and p_i is the purity in leaf i. The lower S is, the better the separation between signal and background. To train a decision tree, one starts with the full training event sample and tests different cut values for each event variable. The combination

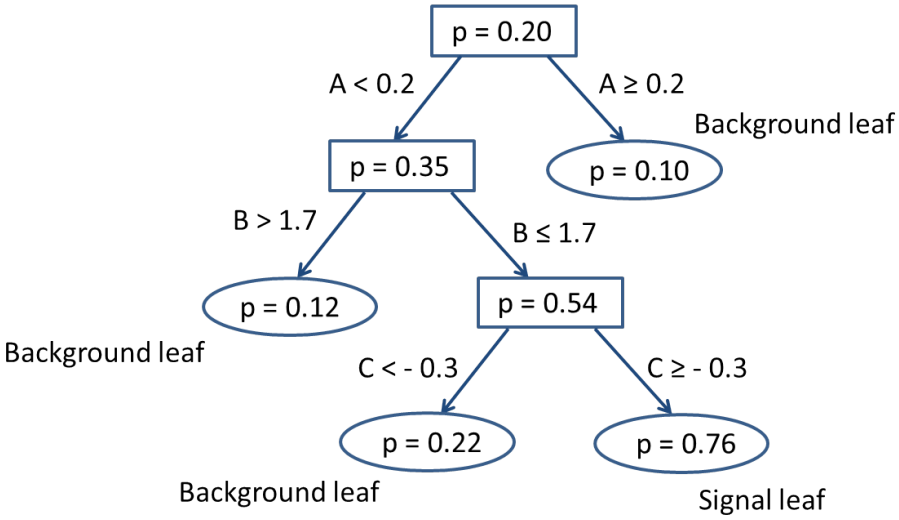


Figure 4.5.: An example of a decision tree with three split nodes (the rectangles) and four leaf nodes (the ellipses): three background leaf nodes and one signal leaf node. The event variables that are used are called A, B and C and at each node the signal purity p is given.

of event variable and cut value which minimizes S is chosen. Each of the two subsamples is then split again following the same rules, and this continues until one of the predefined ending conditions is reached. When that happens, the branch ends in a leaf node. Once all branches end in leaf nodes, the tree is finished and can be used to classify unknown events, resulting in the predicate 'signal' (usually denoted by a tree score of 1) or 'background' (usually denoted by a tree score of -1).

Common ending conditions are that the node contains only background or only signal events, the tree has reached a maximum tree depth, or the best split would result in a node with less than a minimum number of events. The latter two are implemented because otherwise the tree would continue until each node only contains signal or background events. At this point the tree is no longer using general properties of signal or background events, but training on the specific events that are in the training sample, and that means that the decision tree will not perform well on new events. This is called overtraining, and is discussed in more detail in section 4.3.3.

4.3.2. Decision tree forests

Using one decision tree is not very sophisticated, as in general either many events will be wrongly classified (if the tree is not very deep) or there is overtraining (if the tree is deep). To improve the performance, decision tree forests [206] are used, which consist of a large number of decision trees. For most events, some of these trees will classify the event as signal while others classify it as background. The output of the decision tree forest as a whole is the average of the scores of all single trees, possibly taking into account weights (if some of the trees are known or assumed to perform better than others). This output will be a number between 1 and -1, with scores close to 1 for events that are very signal-like and scores close to -1 for events that are very background-like.

The method for training decision trees that was discussed above will always lead to the same tree. There are two methods which are commonly used to produce large numbers of different trees. One is to not always use the same events in the training sample, but randomly take a certain percentage (for example 50%) of the events. Another is to change the weights of the events, specifically to give events which were misclassified a higher weight in the next tree, which improves the classification of these events. This is called boosting, and in the analysis described in this thesis the weights are adjusted using the AdaBoost algorithm [207]. To use this, one calculates

$$\alpha = \beta \cdot \ln\left(\frac{1-f}{f}\right) \tag{4.16}$$

for the tree, with f the fraction of all events that is wrongly classified and β a user-defined boosting strength between 0 and 1. Each event that is wrongly classified is re-weighted with a factor e^{α}, so in real-life situations (where $f < 0.5$) wrongly classified events have a higher weight in the next tree. After this tree has been trained, the boosting is calculated again and new weights are applied, where the boosting is cumulative. Since α is large if the fraction of wrongly classified events is low, it makes sense to use α as the weight of each tree score when calculating the forest score for an event (after all trees have been trained).

In the analysis described in this thesis, decision tree forests which are both boosted and randomized are used, and they are referred to as BDTs.

4.3.3. Overtraining and pruning

One of the largest problems for BDTs is overtraining, which has already been mentioned briefly. One could imagine training a tree such that each leaf contains only signal events or only background events, meaning that the separation is optimal (in this case S would be 0). However, if this tree would be applied to events which are not in the training sample, it would perform worse, because it has been trained to the specifics of the training sample. As the goal is to get a BDT that performs well on unknown samples, this is a problem, and it is referred to as training sample overtraining.

Training sample overtraining can be tested by checking how the BDT performs on new signal and background samples, which are usually referred to as the testing samples (as opposed to the training samples). If there is no overtraining, the BDT score distributions should be the same for the training and testing samples, apart from statistical fluctuations. To quantify how similar the distributions are, the Kolmogorov-Smirnov test is used [208, 209]. The outcome of this test is the probability that the BDT score distributions of the testing and training samples would be as different as they are, or more different, if the two samples were random subsamples of one sample.

There is also a second type of overtraining, which is specific to the situation described in this thesis: the signal sample is simulated while the background sample consists of experimental data. As simulation will never exactly reproduce data, it is possible that the BDT is partially trained to separate data and simulation instead of separating signal and background. In that case, signal data events would be rejected. This is called data-MC overtraining, and it can be tested for by comparing data to background simulation, in this case simulated atmospheric muons and atmospheric neutrinos. The BDT score distributions for data and background simulation should match reasonably well. If for high BDT scores (signal-like

events) the data rates are significantly lower than the rates from simulation, this is a sign of data-MC overtraining. An example of a BDT with significant data-MC overtraining is shown in figure 4.6.

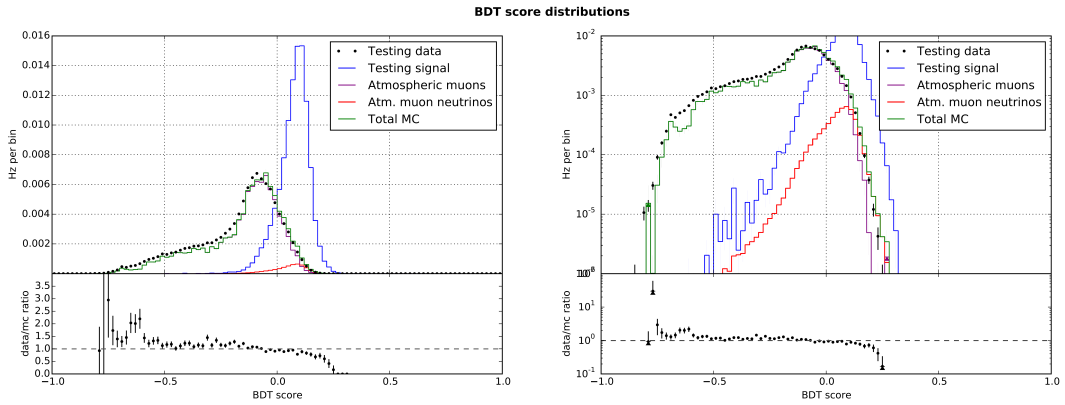


Figure 4.6.: An example of a BDT score distribution for a BDT with significant data-MC overtraining. Data is shown in black, atmospheric muon simulation in purple, atmospheric neutrino simulation in green and the sum of the latter two in blue. Both plots contain the same data, but the left plot has a linear y-scale and the right plot a logarithmic y-scale. For high values of the BDT score, there is significantly less data than simulation. There is also a difference for low BDT scores, but this is related to problems with the simulation, not with the BDT.

Since both types of overtraining discussed above arise from the same issue (the BDT is trained too much on the specifics of the training samples instead of general differences between signal and background), the solutions are also the same. First, one needs to make sure that the statistics in both signal and background samples are large. Second, the randomization discussed in section 4.3.2 helps against overtraining, as each tree is trained under different conditions. Third, it is important that the trees do not go too deep, so that the classification is still based on general properties and not specific events. This can be done by tuning the ending conditions mentioned in section 4.3.1, for example by setting a maximum depth. Finally, a pruning mechanism can be applied to avoid trees that go too deep.

Pruning is applied to each tree separately, after the training of that tree has been finished. The underlying idea is that some splits, especially the later ones, contribute only slightly to the overall separation of signal and background, but have a high risk of contributing to overtraining. Several methods have been devised to identify and remove these splits, to simplify the tree. In the analysis described in this thesis, first same leaf pruning is applied, and then cost complexity pruning.

For same leaf pruning, all split nodes which end in two leaf nodes are considered. If both of these leaf nodes are of the same type (both are background leaves or both are signal leaves), the tree is pruned at that split node (the split node becomes a leaf node). This has no effect on the outcome of the tree and does not contribute to removing overtraining, but it simplifies the tree.

Afterwards, cost complexity pruning is applied. For this type of pruning, the split node which contributes least to the overall separation is determined and removed. This is done

recursively until a predefined percentage of the split nodes has been removed. The node which contributes least to the overall separation is determined by finding the node for which $\Delta S/(n_{leaves} - 1)$ is lowest, where ΔS is the improvement in the Gini impurity (defined in equation 4.15) as a result of this split and n_{leaves} is the total number of leaves in the subtree below this split.

4.3.4. Optimization

The first step in finding the optimal BDT is to choose which event variables are given as input to the BDT. For this analysis, a number of variables were selected by eye (comparing histograms for background and signal) and then tested. In some cases, variables were removed because they caused overtraining (this happened for variables with large differences between distributions for data and simulated background events). In other cases they were removed because they were almost never used in the BDT or because they were highly correlated to other variables. For the BDTs used in the analysis described in this thesis, this process resulted in 9 or 10 variables, depending on the specific BDT. These variables will be discussed in sections 4.4.5 and 4.4.6.

In addition to choosing the event variables to be used, there are also a number of BDT parameters which can be tuned to get an optimal performance, which is defined as the best separation between signal and background while avoiding overtraining. For this analysis, separation was quantified by placing the cut such that there is a fixed background rate, and maximizing the signal efficiency. Overtraining was quantified by comparing data and background simulation for the bins in the BDT score distribution that contain more simulated atmospheric neutrino events than simulated atmospheric muon events. At most one of these bins is allowed to have a difference between data and simulation (with data smaller than simulation) larger than one standard deviation.

The BDT parameters that were tuned for the BDTs described below were the number of trees in the forest (300 to 600), the depth of the trees (4 to 8 cuts), the percentage of split nodes that was removed during pruning (25% to 75%), the boosting strength β (0.25 to 0.75) and the fraction of events that was randomly used for each tree (30% to 100%).

4.4. Event selection

The goal of the event selection is to remove as many background (atmospheric muon) events as possible while keeping as many signal events as possible. In this analysis, there are many different signal samples: as discussed in section 2.5, 19 WIMP masses and 5 annihilation channels are considered. Also, 7 targets at different zeniths (ranging from 102 to 151 degrees) are studied, meaning that signal events range from almost horizontal to close to vertically upgoing.

To optimize the event selection for all these different signal samples, four benchmark samples were considered. These were 100 TeV WIMPs annihilating to muons (a high-energy sample), 100 GeV WIMPs annihilating to b quarks (a low-energy sample), 200 GeV WIMPs annihilating to muons and 500 GeV WIMPs annihilating to b quarks (both in the middle). Also, to avoid optimizing for one specific target, all upgoing events in the signal sample as

described in section 3.4.2 were considered to be signal events.

To make sure that the event selection was optimized for the most useful events, it was optimized on a 'golden sample' of well-reconstructed Monte Carlo signal events. This golden sample was used to calculate efficiencies, make histograms and determine cuts, while the final analysis is done on the full Monte Carlo sample. The golden sample contains only charged-current muon neutrino events, with a reconstruction error of less than 5 degrees (for the samples with annihilation to muons) or 10 degrees (for the samples with annihilation to b quarks). The difference is due to the fact that the (low-energy) samples with annihilation to b quarks are less well-reconstructed, so requiring a reconstruction error of less than 5 degrees would result in a very small sample. Also, the golden sample contains only events where at least six DOMs on at least two strings were hit. This is because at least six measurements are necessary to reconstruct the track, and if all hits are on one string, the azimuth cannot be reconstructed due to the rotational symmetry.

The event selection consists of several different levels with different goals. Throughout this section, the effect of each level on the different simulated event samples and a test data sample will be shown by giving the rates after the cuts and calculating the fraction of the samples that passes these cuts. For this, the following are used: an atmospheric muon and atmospheric muon neutrino sample (the generation of which was discussed in section 3.4), the signal samples discussed in the paragraph above, and a small subsample of the data (approximately 10%).

Both the data subsample and the simulation samples used for designing the event selection are from the first year out of the three years of data that are being used for the analysis. There are a number of changes in the first steps of the processing, which will be mentioned where necessary, but as will be shown in section 4.5, at the last level of the event selection the three years are compatible within statistical fluctuations.

4.4.1. Filters

There are several filters in place at the South Pole which are applied to events which pass the triggers. Most of these select specific event types, for example cascades or low-energy events, which means that they can be used as a first step to select interesting events. For the analysis described in this thesis, interesting events are charged-current muon neutrino events (often referred to as 'tracks') with energies ranging from about 10 GeV (the energy threshold of DeepCore) to 100 TeV. The three filters which best match this are the Muon, DeepCore and LowUp filters. The first level of the event selection removes events which do not pass any of these filters.

DeepCore filter

The DeepCore filter has been designed to select events where (almost) all of the light is produced inside DeepCore, the low-energy subdetector located at the bottom middle of IceCube (see section 3.3.1). These events are all low-energy events, since a 100 GeV muon already has a range of approximately 500 meters, which is larger than the size of DeepCore. The DeepCore filter is only applied to events which pass the SMT3 trigger described in section 3.3.3.

The filter, which is described in [176], was based on all HLC hits for IC86-2011. For IC86-2012 and IC86-2013, it used all hits after the first step of the hit cleaning described in section 4.1.1, which also includes part of the SLC hits. First, the mean and the standard deviation of the times of these hits in the DeepCore fiducial volume are determined. Then, all hits with a time more than one standard deviation from the mean are removed. The remaining hits are used to calculate the center of gravity (CoG), which is the charge-weighted average location of these hits. Then, the time of each hit is corrected for the time it would take unscattered light to travel from the CoG to the location of the hit. Finally, the average of the corrected times is calculated.

Once the time and location of the CoG have been reconstructed as discussed above, all hits in the hit series (HLC hits or cleaned hits) outside the DeepCore fiducial volume are considered. For each hit, the speed that a particle would have if it travelled from this hit to the CoG is calculated. If both the hit outside DeepCore and the hits in DeepCore were caused by one muon, this speed would be approximately 0.3 m/ns. If the speed is between 0.25 and 0.4 m/ns (taking into account that a possible muon does not travel directly from the location of one hit to the location of another), the hit is tagged as a veto hit. Only events which have less than two veto hits pass the DeepCore filter.

The average rate of the DeepCore filter was approximately 27 Hz in IC86-2011, and approximately 17 Hz in IC86-2012 and IC86-2013. The lower rate was mainly due to the fact that more hits were used for vetoing, so there was a higher probability that atmospheric muons were vetoed.

Muon filter

The goal of the Muon filter is to select tracks (muons) which have a relatively high probability of being due to neutrino interactions. It selects both upgoing and downgoing tracks, as there are also interesting possible neutrino sources in the Southern Hemisphere. However, for downgoing tracks the cuts (for example on the total charge deposited in the detector) are significantly harder, to avoid selecting too many atmospheric muon events.

The Muon filter is applied to all events that pass the SMT8 trigger described in section 3.3.3 and uses the zenith from the PandelSPE reconstruction, $\theta_{PandelSPE}$. In addition to this, two cut variables are used: for events reconstructed as upgoing, $\text{rlogl3}_{PandelSPE}$ (defined in equation 4.6) is used to remove badly reconstructed tracks (which are likely to be downgoing). For events reconstructed as downgoing, the sum of the charges of all hits Q_{tot} is used to remove low-energy tracks, which are more likely to be due to atmospheric muons. The charge is expressed in units of PE, which is the most likely deposited charge for a single photon (see section 3.3.2).

Specifically, the cuts are as follows:

$$\begin{aligned}
\theta_{PandelSPE} \leq 60^\circ \quad &\& \quad \log(Q_{tot}) \geq 0.6 \cdot \cos(\theta_{PandelSPE}) + 2.6 \text{ or} \\
60^\circ < \theta_{PandelSPE} \leq 78.5^\circ \quad &\& \quad \log(Q_{tot}) \geq 3.9 \cdot \cos(\theta_{PandelSPE}) + 2.6 \text{ or} \\
\theta_{PandelSPE} > 78.5^\circ \quad &\& \quad \text{rlogl3}_{PandelSPE} \leq 8.7
\end{aligned}$$

For each of the three years considered in this analysis, the average rate of the Muon Filter was approximately 38 Hz.

LowUp filter

The final filter used in this analysis is the LowUp filter, which selects low-energy upgoing events. It is applied to events which pass at least one of the four triggers described in 3.3.3. It consists of cuts on a number of variables, some of which ($\theta_{PandelSPE}$ and z_{trav}, both measures for how upgoing an event is) have been discussed before, while some others are new: N_{ch} (the total number of DOMs with at least one hit), z_{ext} (the distance in z between the hits with the highest and lowest z values), t_{ext} (the time between the first and last hit), N_{in} (the number of hits on strings that are not on the boundary of the detector) and N_{up5} (the number of hits in DOMs on the upper five layers of the detector). The cuts that are applied are:

$$
\begin{aligned}
z_{ext} &< 600 \text{ m} \\
t_{ext} &< 4000 \text{ ns} \\
N_{ch} &> 3 \text{ (IC86-2011 and IC86-2012) or } N_{ch} > 4 \text{ (IC86-2013)} \\
\theta_{PandelSPE} &> 80^\circ \\
z_{trav} &> -10 \text{ m} \\
N_{in} &> 0 \\
N_{up5} &= 0
\end{aligned}
$$

The resulting rate was approximately 30 Hz for each of the three years considered in this analysis.

Combined rates

As there is overlap between the LowUp filter and the other filters (since some events fulfill the criteria of multiple filters), the total rate at filter level is lower than the sum of the filter rates mentioned so far. For IC86-2011, it was approximately 92 Hz and for IC86-2012 and IC86-2013 it was approximately 81 Hz.

In table 4.1, the rates at filter level are shown for the test data, atmospheric muon and atmospheric muon neutrino samples. Since the test data is only a small subset of all experimental data for 2011, the rate is different from the average rate mentioned above. Since the rate of signal neutrinos (if they exist) depends on the annihilation cross section, which is unknown, it is not possible to give absolute rates for these. From the table it is clear that at filter level, the data is highly dominated by atmospheric muons, and also that data and simulation do not match well: there are too few simulated atmospheric muons that pass these filters, compared to the actual data. This is partially due to the fact that some of the events passing the filters are due to noise hits only and not to a particle passing through the detector, and partially due to issues with the atmospheric muon simulation. As mentioned before, at this level the discrepancy between data and simulation is not a problem for this analysis, since the atmospheric muon simulation is not used in designing the event selection or in the final analysis.

Sample	Rate [Hz]
Data	85.95
Atmospheric μ simulation	69.57
Atmospheric ν_μ simulation	0.0184

Table 4.1.: The rates of events passing at least one of the three filters described in section 4.4.1, for test data, atmospheric muon simulation and atmospheric muon neutrino simulation. The difference between the total rates for data and simulation is partially due to the lack of pure noise simulation (events where no particle passes through the detector, but the detector is triggered by the combination of several noise hits) and partially due to issues with the atmospheric muon simulation.

4.4.2. Good run selection

The good run selection is different from the other steps in the event selection because it does not remove single events, but full runs. IceCube data is usually taken in runs of approximately 8 hours, but not all runs can be used for the data analysis. First, all runs which did not have the standard configuration (for example because they were used for calibration, like the flasher runs discussed in section 3.3.4) are removed. Then, all runs for which at least one of the strings was not active are removed, as even one (veto) string not being active can have a significant effect on for example the effectiveness of the DeepCore filter. Finally, the stability of the event rate at filter level is tested by comparing the event rate for each run with the average of the five 'good' runs (runs passing all criteria) before this run. If the deviation is larger than 5%, the run is also removed.

The filter rates for the three filters being used in this analysis (the Muon filter, the LowUp filter and the DeepCore filter) over the three years of data used in this analysis are shown in figure 4.7. Here, the last step in the good run selection has not been applied yet, and the stars denote runs which have deviations larger than 5% for at least one filter.

The largest visible feature is the approximately sinusoidal behaviour of the rates for all filters. This is due to seasonal variations: if the air above the detector is colder (in the Southern winter), the atmosphere is denser. That means that the probability that pions and kaons interact before decaying is larger, which means that fewer atmospheric muons will reach the detector. There are also features at the changes between data-taking seasons (also referred to as 'years'): the IC86-2012 DeepCore filter rate is much lower than the IC86-2011 rate because the filter changed significantly, and the rates of the other filters also changed because of changes in calibration. The IC86-2013 LowUp filter rate is lower than that of the previous years because the cut on the number of hit DOMs was stricter. Finally, there is a change in the filter rates at run number 118495 which is due to the correction of a bug which caused the deadtime of the DOMs to be too large.

After the good run selection is applied, the IC86-2011 season consists of 1115 runs (333.4 days), the IC86-2012 season consists of 1116 runs (329.1 days) and the IC86-2013 season consists of 1311 runs (351.4 days). Since part of the IC86-2011 data is used in designing the event selection (for example for training and testing the BDTs), these runs are removed from the final analysis sample to avoid biases, and 1004 runs (302.2 days) remain in the IC86-2011 sample used for the final analysis.

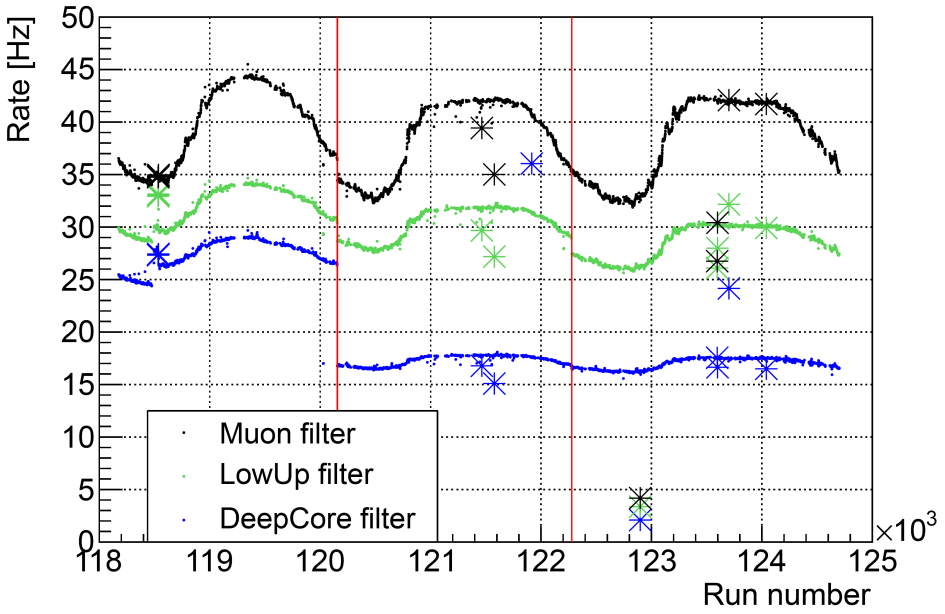


Figure 4.7.: The filter rates for the three filters being used in this analysis as a function of run number for all three years of data that are being used (IC86-2011 to IC86-2013, the x axis runs from May 2011 to May 2014). The red vertical lines denote when the switch to a new year of data-taking took place. The stars denote runs which have deviations larger than 5% for at least one filter, these runs are not used for the analysis.

4.4.3. Preliminary cuts

The event selection described up until now is a general event selection used for many IceCube analyses, but from this section onwards the event selection described was developed specifically for this analysis. The first step of this part consists of straightforward preliminary cuts. The goal of these is to significantly reduce the background by selecting upgoing, well-reconstructed events which are not due to noise. As mentioned in the description of the golden sample at the beginning of section 4.4, at least six hit DOMs on at least two strings are necessary to be able to reconstruct the direction, so the first cut is to require that the number of hit DOMs $N_{ch} > 5$ and the number of hit strings $N_{str} > 1$.

To remove events which are reconstructed as downgoing, $\theta_{PandelSPE} > 80$ degrees or $\theta_{PandelMPE} > 80$ degrees is required (where θ_x is the zenith as reconstructed by reconstruction x). To remove events which are badly reconstructed, rlogl2.5$_{PandelSPE} < 8.5$ or rlogl2.5$_{PandelMPE} < 8.5$ is required (rlogl was defined in equation 4.6).

Finally, a significant fraction of the events (estimated to be a few Hz [210]) that pass the DeepCore filter consist purely of noise hits. To remove these events, one can use the fact that noise triggers will consist of a low number of hits and there will be no preferred direction in the event. An algorithm is applied to events with less than 20 hits which pass the DeepCore filter (all other events automatically pass this cut). The first step is to determine which 750 ns wide time window encompasses the most hits. Then, all hit pairs which are in this time window and have an apparent velocity (the distance between two hits divided by their time difference) between 0.1 and 1.0 m/ns are binned in 48 directional bins. If there is at least one bin with 3 or more hit pairs (which implies that there is a preferred direction in the event),

Sample	Rate after cuts [Hz]	Passing fraction of cuts
Data	17.77	20.7%
Atmospheric μ simulation	16.99	24.4%
Atmospheric ν_μ simulation	0.0125	67.7%
100 TeV $\mu\mu$ signal simulation		99.8%
200 GeV $\mu\mu$ signal simulation		94.7%
500 GeV bb signal simulation		92.4%
100 GeV bb signal simulation		86.4%

Table 4.2.: The rates and passing fractions for events that pass the preliminary cuts described in section 4.4.3 for test data, atmospheric muon simulation, atmospheric muon neutrino simulation and four signal samples, where bb or $\mu\mu$ refers to the annihilation channel and the mass refers to the mass of the WIMP.

the event passes the cut.

In table 4.2, the effect of the cuts described above on the different (data and simulation) samples is shown. These cuts greatly improve the agreement in rates between data and simulation, and for a large part this is due to the noise removal algorithm discussed above (the simulation samples do not contain noise-triggered events).

4.4.4. IceCube-dominated and DeepCore-dominated subsamples

The energy range of the signal neutrinos that are searched for in this analysis is large, from 10 GeV up to 100 TeV. This leads to different event signatures in the detector. To get an event selection that works well for all WIMP masses and annihilation channels (which lead to different energy spectra), the event sample is split up in two subsamples, and different cuts are applied to both.

The low-energy subsample (which is referred to as the DeepCore-dominated sample) contains all events where the number of hit DOMs in the DeepCore fiducial region (defined in section 3.3.1) is larger than the number of hit DOMs outside that region, and where the number of hit DOMs outside that region is smaller than 8. The latter requirement is there to make sure that high-energy events which happen to pass through DeepCore are in the high-energy sample. The high-energy subsample, which is referred to as the IceCube-dominated sample, contains all events which are not part of the DeepCore-dominated sample.

Figure 4.8 shows which percentage of simulated atmospheric muon neutrino events that pass the preliminary cuts ends up in which sample, as a function of true neutrino energy. As expected, for low energies (up to 50 GeV) most of the events which pass the preliminary cuts are in the DeepCore-dominated sample, while for higher energies (above approximately 200 GeV) almost all events are in the IceCube-dominated sample. Table 4.3 shows which percentage of the various data and simulation samples ends up in which subsample.

4.4.5. Second level of cuts

As discussed in section 4.1.2, the goal is to eventually use SplineMPE as track reconstruction for this analysis, as it performs best. However, it is not feasible to apply SplineMPE to all events after the preliminary cuts, since this would take too much processing time. For

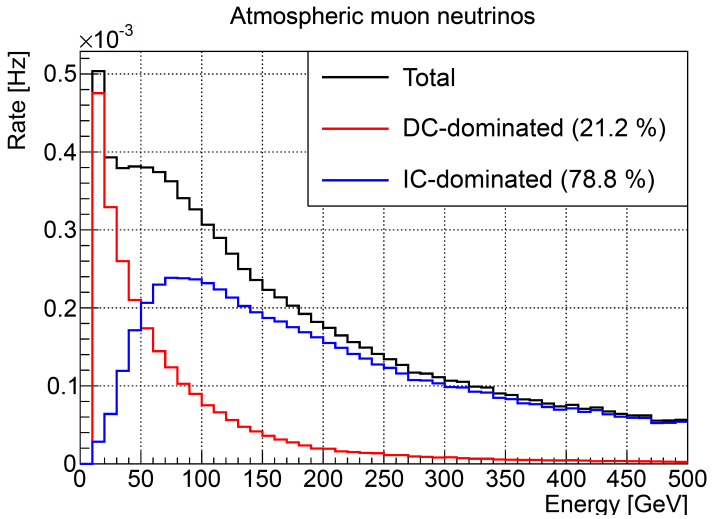


Figure 4.8.: The rates of events which end up in the DeepCore-dominated or the IceCube-dominated sample as a function of true neutrino energy, using a simulated atmospheric muon neutrino sample.

Sample	IceCube-dominated	DeepCore-dominated
Data	97.4% (17.32 Hz)	2.6% (0.459 Hz)
Atmospheric μ simulation	97.0% (16.47 Hz)	3.0% (0.517 Hz)
Atmospheric ν_μ simulation	78.8% (9.81 mHz)	21.2% (2.64 mHz)
100 TeV $\mu\mu$ signal simulation	99.5%	0.5%
200 GeV $\mu\mu$ signal simulation	79.1%	20.9%
500 GeV bb signal simulation	69.8%	30.2%
100 GeV bb signal simulation	25.7%	74.3%

Table 4.3.: The percentage of events that ends up in the DeepCore-dominated or IceCube-dominated subsamples (described in section 4.4.4) for test data, atmospheric muon simulation, atmospheric muon neutrino simulation and four signal samples, where bb or $\mu\mu$ refers to the annihilation channel and the mass refers to the mass of the WIMP.

this reason, the cut level described in this section was designed to reduce the data rate to approximately 100 mHz for each of the subsamples, at which level it is possible to apply the more time-consuming reconstructions. As is shown in table 4.3, for the DeepCore-dominated sample the experimental data rate is already relatively close to this, so a few straightforward cuts are enough. For the IceCube-dominated sample, the rate needs to be reduced by a much larger factor, and this is done by using a cut on the output of a BDT.

DeepCore-dominated sample

As is shown in table 4.3, after the preliminary cuts the DeepCore-dominated sample is less dominated by atmospheric muons than the IceCube-dominated sample. This is due to the fact that if an atmospheric muon passes through DeepCore, it must first have passed through the whole IceCube detector and the probability that it only created 7 or less hits there is low. However, the DeepCore-dominated sample is still dominated by dim atmospheric muons, which only create a few hits outside DeepCore. To remove a significant fraction of these, at this level two variables are used. The first one is z_{trav}, which is a measure for how upgoing

Sample	Rate after cuts [mHz]	Passing fraction of cuts
Data	128.3	28.0%
Atmospheric μ simulation	150.3	29.1%
Atmospheric ν_μ simulation	2.56	96.9%
500 GeV bb signal simulation		96.3%
100 GeV bb signal simulation		98.9%

Table 4.4.: The rates after and passing fractions for the second level of cuts described in this section for the DeepCore-dominated sample for test data, atmospheric muon simulation, atmospheric muon neutrino simulation and two signal samples, where bb refers to the annihilation channel and the mass refers to the mass of the WIMP.

an event is, and which was discussed in section 4.2.1.

The second variable is the result of an algorithm which was specifically designed to remove dim atmospheric muons which cause only a few hits in the detector. Often, these hits are far enough apart that they are removed by the cleaning described in section 4.1.1. For this algorithm, the original hit series (before cleaning) is used. All hits outside the DeepCore fiducial volume which were detected in the 5 μs before the time of the first trigger in the event are considered. Each of these hits is treated as the seed of a cluster: if there are any hits within 250 meters and 1 μs of a hit in the cluster, those are added to the cluster, and this is done recursively until no more hits can be added. The total charge of the largest cluster (in units of PE) is called Q_{RT}, and this is the second variable that is used at this level.

Histograms of both variables are shown in figures 4.9 and 4.10. The cuts that are applied, which are also shown in the figures, are $Q_{RT} < 6$ PE and $z_{trav} > -40$ m. As can be seen from the figures, these are loose cuts that remove only a small fraction of the simulated signal, while significantly reducing the data rate. This is quantified in table 4.4, where the high-energy signal samples are no longer shown as the event selection for the DeepCore-dominated sample is optimized for the low-energy signal samples only.

IceCube-dominated sample

From table 4.3 one can see that after the preliminary cuts, the IceCube-dominated subsample still has a data rate of 17.3 Hz, so to get down to 100 mHz approximately 99.4% of the data events need to be removed. For this, a BDT was shown to perform significantly better than straight cuts. Such a BDT needs to be trained on a particular signal sample. In order to train a BDT that performs well for different signal samples (with different energy distributions), the signal sample consisted of the average of the spectra of 100 TeV WIMPs, 10 TeV WIMPs, 1 TeV WIMPs and 200 GeV WIMPS, all annihilating to muons.

The BDT was optimized following the procedure described in section 4.3.4. During this optimization it was found that while the speed as reconstructed by LineFit (v_{LF}, see section 4.1.2) is a valuable variable, the wide range in values was problematic. For this reason, a precut of $v_{LF} < 2$ m/ns was used (which removes only badly reconstructed events, as a relativistic muon has a velocity of 0.3 m/ns) before training the BDT.

The optimization process resulted in the following nine variables to be used in the BDT:

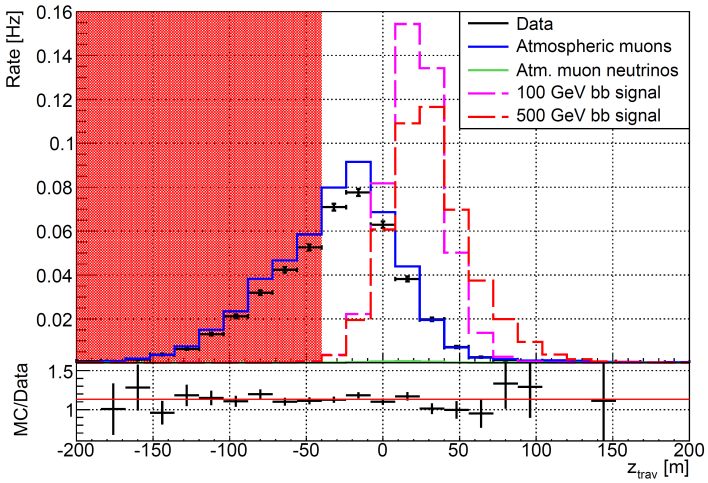


Figure 4.9.: Histogram of z_{trav} for the DeepCore-dominated subsample after the preliminary cuts, for a subset of the experimental data, simulated atmospheric muons, simulated atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs annihilating to b quarks and 100 GeV WIMPs annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown. Events in the red area ($z_{trav} < -40$ m) are removed.

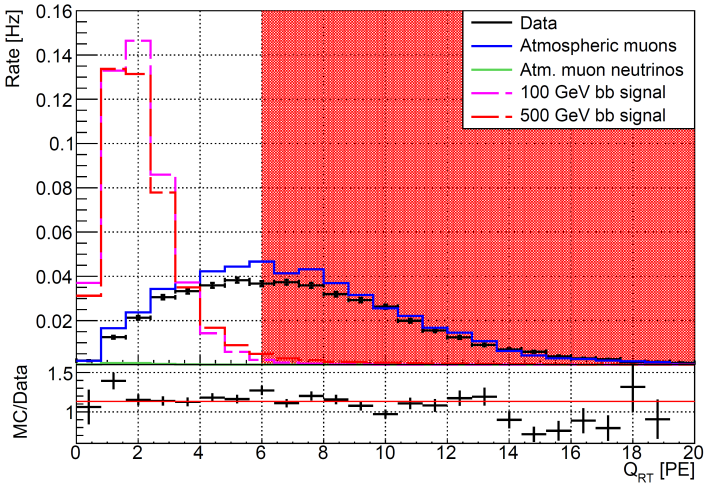


Figure 4.10.: Histogram of Q_{RT} for the DeepCore-dominated subsample after the preliminary cuts, for a subset of the experimental data, simulated atmospheric muons, simulated atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs annihilating to b quarks and 100 GeV WIMPs annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown. Events in the red area ($Q_{RT} > 6$ PE) are removed.

- z_{trav}, the average drift in the z direction with respect to the first quarter of hits (see section 4.2.1),
- L_{dir}, the track length along which there are direct hits, for the PandelMPE reconstruction (see section 4.2.3),
- N_{dir}, the number of direct hits, for the PandelMPE reconstruction (see section 4.2.3),
- θ_{LF}, the zenith as reconstructed by LineFit (see section 4.1.2),
- v_{LF}, the speed as reconstructed by LineFit (see section 4.1.2),
- rlogl2.5, the reduced log-likelihood for the PandelMPE reconstruction (see section 4.1.2),
- $\langle d_{ht} \rangle$, the charge-weighted average track-DOM distance, for the PandelMPE reconstruction (see section 4.2.4),
- $\Delta\Psi$(LineFit, PandelMPE), the angle between the tracks as reconstructed by LineFit and PandelMPE (see section 4.2.4),
- N_{sub}, the number of subevents found by the Topological Splitting algorithm (see section 4.2.5).

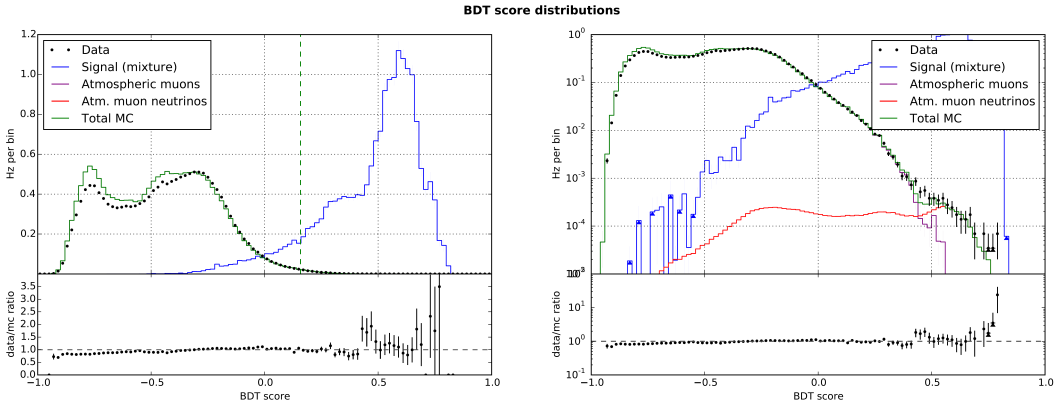


Figure 4.11.: Histogram of the BDT score for the IceCube-dominated subsample after the preliminary cuts, for a subset of the experimental data, simulated atmospheric muons, simulated atmospheric muon neutrinos, and the signal sample (consisting of a normalized average of the spectra for 100 TeV, 10 TeV, 1 TeV and 200 GeV WIMPs annihilating to muons). For both experimental data and signal, the testing samples are shown, not the training samples. Both plots contain the same data, but the left plot has a linear y-scale and the right plot a logarithmic y-scale. Below the main plot, the ratio of Monte Carlo background simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown. The green dashed line shows the cut that is applied, all events with BDT scores lower than this cut value are removed.

Histograms for data, signal simulation (normalized to the data) and background simulation for these variables are shown in appendix A. The resulting BDT score distribution for the different samples is shown in figure 4.11. This figure shows that there is a separation between atmospheric muons (the reducible background) and atmospheric or signal neutrinos, with atmospheric muons on average getting low BDT scores while neutrinos get high BDT scores. Also, in the most interesting region (the signal-like region), there is reasonable agreement

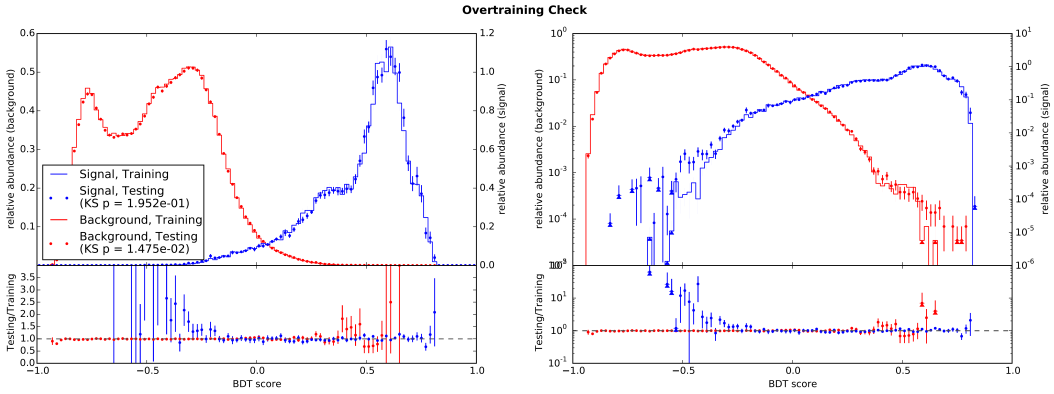


Figure 4.12.: Comparison of histograms of the BDT score for the background and signal testing and training subsamples, for the BDT trained on the IceCube-dominated subsample after the preliminary cuts. Both plots contain the same data, but the left plot has a linear y-scale and the right plot a logarithmic y-scale. KS p denotes the p-value that is the result of the Kolmogorov-Smirnov test described in section 4.3.3.

Sample	Rate after cuts [mHz]	Passing fraction of cuts
Data	100.0	0.58%
Atmospheric μ simulation	79.6	0.48%
Atmospheric ν_μ simulation	4.91	46.6%
200 GeV $\mu\mu$ signal simulation		78.3%
100 TeV $\mu\mu$ signal simulation		96.2%

Table 4.5.: The rates and passing fractions for events in the IceCube-dominated sample that pass the cuts described in this section for test data, atmospheric muon simulation, atmospheric muon neutrino simulation and two signal samples, where $\mu\mu$ refers to the annihilation channel and the mass refers to the mass of the WIMP.

between data and simulation, and there are no signs of data-MC overtraining. In figure 4.12, the BDT score distributions for the training and testing samples are compared, and the resulting p-value of the Kolmogorov-Smirnov test is given. There is no sign of training sample overtraining. From this it can be concluded that the BDT performs well.

The goal of this cut level is to reduce the data rate to approximately 0.1 Hz, which is reached if all events with a BDT score below 0.1582 are removed. The dashed green line in figure 4.11 shows where this cut value is located. The resulting rates are given in table 4.5, where one can see that after these cuts approximately 5% of the data sample is expected to consist of (atmospheric) neutrinos. In this table, the low-energy signal samples are no longer shown as the event selection for the IceCube-dominated sample is optimized on the high-energy signal samples only.

4.4.6. Final cut level

To get a sample that is dominated by (atmospheric) neutrinos, another cut level is necessary, and for both the IceCube-dominated and the DeepCore-dominated sample a BDT performs better than cuts on single variables here. Since all events in the sample have now been reconstructed by both SplineMPE and BootSPE, it is possible to use variables based on these

reconstructions for the BDT. The final cut on the resulting BDT score will be optimized to get the best sensitivity to dark matter.

DeepCore-dominated sample

For the DeepCore-dominated sample, the signal sample used for training and testing the BDT is that of 500 GeV WIMPs annihilating to b quarks. The BDT was optimized following the procedure described in section 4.3.4, which resulted in the following set of event variables to be used:

- z_{trav}, the average drift in the z direction with respect to the first quarter of hits (see section 4.2.1),
- z_{HLC}, the z position of the first HLC hit (see section 4.2.2),
- l_{HLC}, the string layer of the DOM that detected the first HLC hit (see section 4.2.2),
- $\langle d_{ht} \rangle$, the charge-weighted average track-DOM distance for the SplineMPE reconstruction (see section 4.2.4),
- $\theta_{StrictBootSPE}$, the zenith as reconstructed by BootSPE, using a stricter hit cleaning (see sections 4.1.2 and 4.2.1),
- r_{FR} the r value (where $r = \sqrt{x^2 + y^2}$) of the track starting point as reconstructed by FiniteReco (see section 4.1.2),
- N_{dir}, the number of direct hits, for the SplineMPE reconstruction (see section 4.2.3),
- $\Delta\Psi$(LineFit, SplineMPE), the angle between the tracks as reconstructed by LineFit and SplineMPE (see section 4.2.4),
- θ_{geo}, the smallest zenith (as reconstructed by PandelSPE) of the geometrically split subevents (see section 4.2.5).

Histograms for data, signal simulation (with arbitrary normalization) and background simulation for these variables are shown in appendix B. The resulting BDT score distribution for different samples is shown in figure 4.13, which shows the separation between atmospheric muons and atmospheric (or signal) neutrinos. Again, there is reasonable agreement between data and simulation and there are no signs of data-MC overtraining. In figure 4.14, the BDT score distributions for the training and testing samples are compared, and the result of the Kolmogorov-Smirnov test is given as a p-value. There is no sign of training sample overtraining. From this it can be concluded that the BDT performs well.

IceCube-dominated sample

From figure 4.8 one can see that most of the events with true neutrino energies above approximately 100 GeV end up in the IceCube-dominated sample, which means that the range of signal neutrino energies is large (from 100 GeV to 100 TeV). In general, it is easier to separate high-energy neutrinos from atmospheric muons than to separate low-energy neutrinos from atmospheric muons, as high-energy neutrinos are better reconstructed and there is less background in that energy range. However, many of the WIMP masses that are being tested will only produce relatively low-energy neutrinos, so only optimizing on the high-energy neutrinos

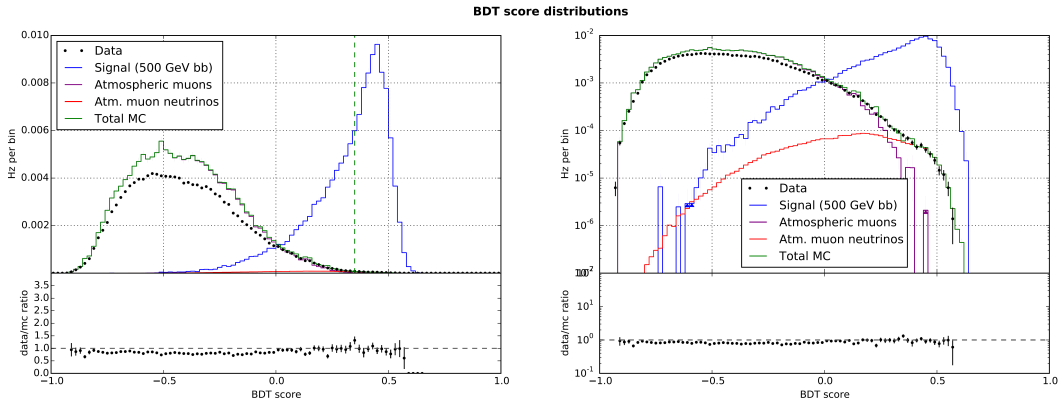


Figure 4.13.: Histogram of the BDT score for the DeepCore-dominated subsample after the second level of cuts, for a subset of the experimental data, simulated atmospheric muons, simulated atmospheric muon neutrinos, and the signal sample (500 GeV WIMPs annihilating to b quarks). For both experimental data and signal, the testing samples are shown, not the training samples. Both plots contain the same data, but the left plot has a linear y-scale and the right plot a logarithmic y-scale. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

is not a solution.

To solve this problem, two BDTs for the IceCube-dominated sample are trained at the final cut level: one on a relatively low-energy signal sample and one on a high-energy signal sample. To determine which of these BDTs performs best for each signal sample that is being tested in the final analysis, the sensitivity using each of the BDTs is calculated. This will be described in more detail in chapter 5. The two BDTs are referred to as the 'soft BDT' (the one trained on the low-energy sample) and the 'hard BDT' (the one trained on the high-energy sample).

The hard BDT is trained on a sample of 10 TeV WIMPs annihilating to muons. The following event variables are used:

- θ_{all}, the smallest zenith (as reconstructed by PandelSPE) of both the geometrically split subevents and the subevents split in time (see section 4.2.5),
- N_{dir}, the number of direct hits, for the BootSPE reconstruction (see section 4.2.3),
- L_{dir}, the track length along which there are direct hits, for the BootSPE reconstruction (see section 4.2.3),
- z_{COG}, the z position of the Center of Gravity (the charge-weighted average location of all hits),
- $\theta_{StrictBootSPE}$, the zenith as reconstructed by BootSPE, using a stricter hit cleaning (see sections 4.1.2 and 4.2.1),
- $\langle d_{ht} \rangle$, the charge-weighted average track-DOM distance for the SplineMPE reconstruction (see section 4.2.4),
- Q_b/Q_{tot}, the fraction of the total charge that is on the boundary layer of the detector (see section 4.2.6),

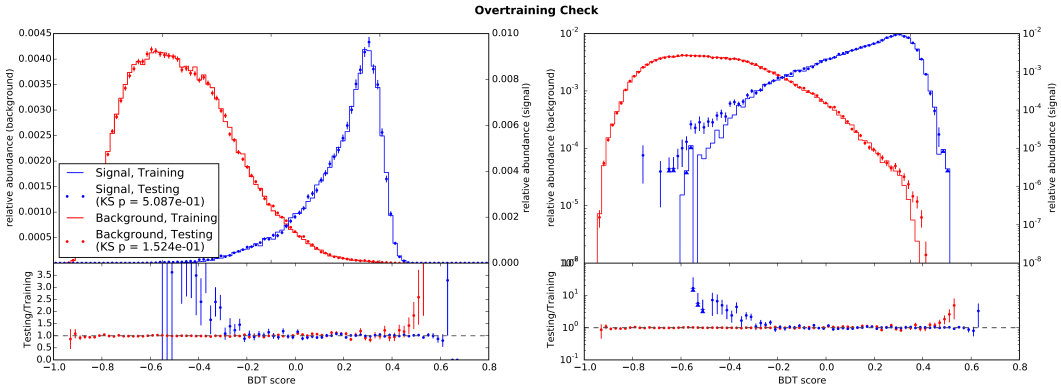


Figure 4.14.: Comparison of histograms of the BDT score for the background and signal testing and training subsamples, for the BDT trained on the DeepCore-dominated subsample after the second level of cuts. Both plots contain the same data, but the left plot has a linear y-scale and the right plot a logarithmic y-scale. KS p denotes the p-value that is the result of the Kolmogorov-Smirnov test described in section 4.3.3.

- $\sigma_{BootSPE}$, the estimated error of the BootSPE reconstruction (see section 4.1.2),
- Q_s/Q_{tot}, the fraction of the total charge that is on the string with the highest charge (see section 4.2.6),
- z_{patt}, the z pattern (see section 4.2.1).

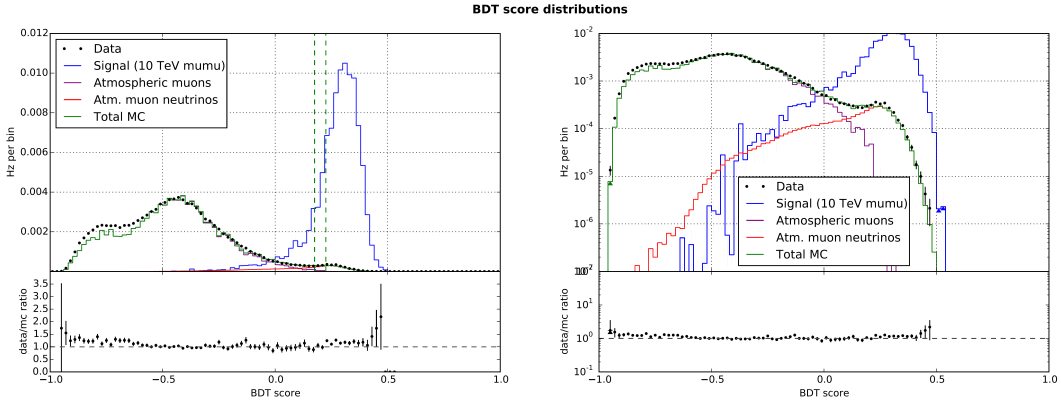


Figure 4.15.: Histogram of the 'hard' BDT score for the IceCube-dominated subsample after the second level of cuts, for a subset of the experimental data, simulated atmospheric muons, simulated atmospheric muon neutrinos, and the signal sample (10 TeV WIMPs annihilating to muons). For both experimental data and signal, the testing samples are shown, not the training samples. Both plots contain the same data, but the left plot has a linear y-scale and the right plot a logarithmic y-scale. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

Histograms for data, signal simulation (with arbitrary normalization) and background simulation for these variables are shown in appendix C. The resulting BDT score distribution for different samples is shown in figure 4.15, which shows the separation between atmospheric muons and atmospheric (or signal) neutrinos.

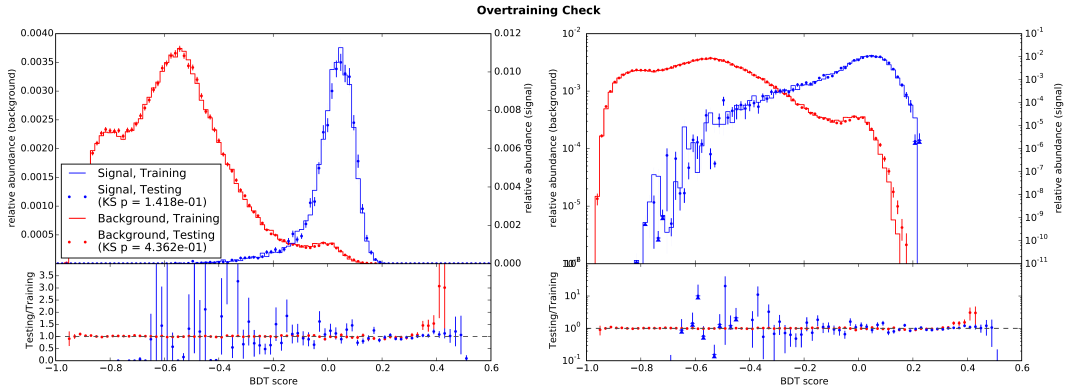


Figure 4.16.: Comparison of histograms of the BDT score for the background and signal testing and training subsamples, for the 'hard' BDT trained on the IceCube-dominated subsample after the second level of cuts. Both plots contain the same data, but the left plot has a linear y-scale and the right plot a logarithmic y-scale. KS p denotes the p-value that is the result of the Kolmogorov-Smirnov test described in section 4.3.3.

The soft BDT is trained on a sample of 500 GeV WIMPs annihilating to muons. The following event variables are used:

- θ_{all}, the smallest zenith (as reconstructed by PandelSPE) of both the geometrically split subevents and the subevents split in time (see section 4.2.5),
- L_{emp}, the maximum length along the track (as reconstructed by BootSPE) with no hit DOMs within a radius of 150 meters (see section 4.2.4),
- $\Delta\Psi$(BootSPE, SplineMPE), the angle between the tracks as reconstructed by BootSPE and SplineMPE (see section 4.2.4),
- $\theta_{StrictBootSPE}$, the zenith as reconstructed by BootSPE, using a stricter hit cleaning (see sections 4.1.2 and 4.2.1),
- $\langle d_{ht} \rangle$, the charge-weighted average track-DOM distance for the SplineMPE reconstruction (see section 4.2.4),
- L_{dir}, the track length along which there are direct hits, for the BootSPE reconstruction (see section 4.2.3),
- z_{HLC}, the z position of the first HLC hit (see section 4.2.2),
- $\sigma_{BootSPE}$, the estimated error of the BootSPE reconstruction (see section 4.1.2),
- Q_s/Q_{tot}, the fraction of the total charge that is on the string with the highest charge (see section 4.2.6),
- z_{patt}, the z pattern (see section 4.2.1).

Histograms for data, signal simulation (with arbitrary normalization) and background simulation for these variables are shown in appendix C. The resulting BDT score distribution (using the testing sample for data and signal simulation) is shown in figure 4.17, which shows the separation between atmospheric muons and atmospheric (or signal) neutrinos.

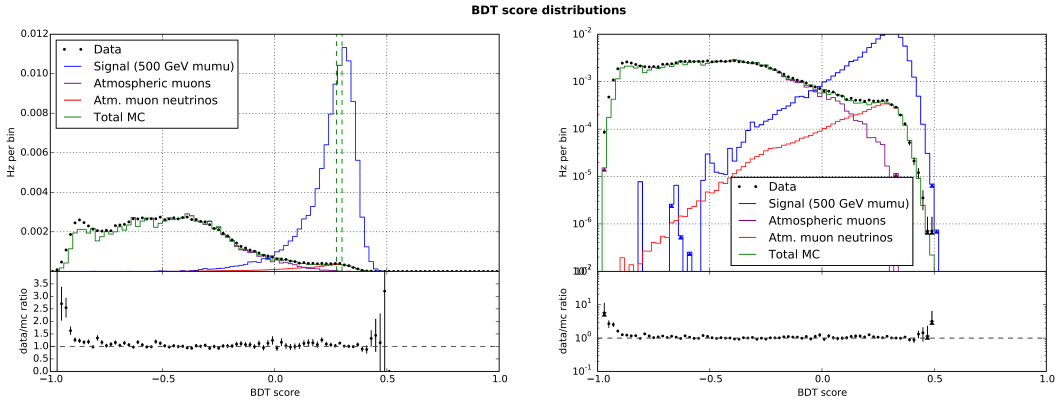


Figure 4.17.: Histogram of the 'soft' BDT score for the IceCube-dominated subsample after the second level of cuts, for a subset of the experimental data, simulated atmospheric muons, simulated atmospheric muon neutrinos, and the signal sample (500 GeV WIMPs annihilating to muons). For both experimental data and signal, the testing samples are shown, not the training samples. Both plots contain the same data, but the left plot has a linear y-scale and the right plot a logarithmic y-scale. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

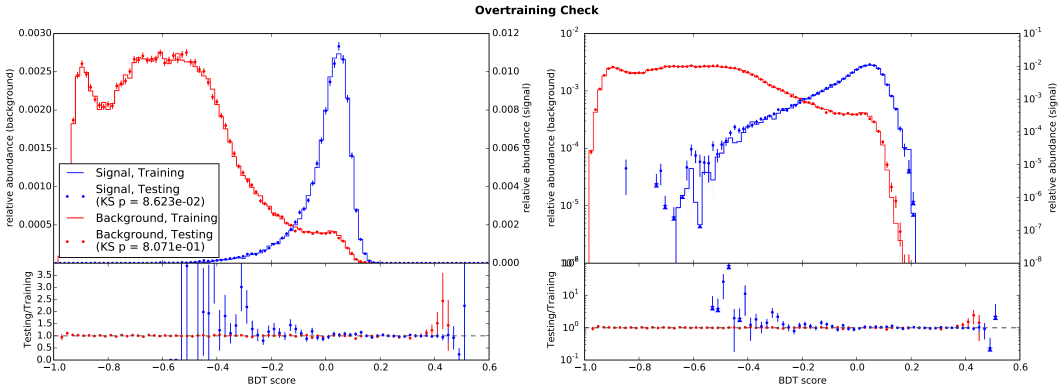


Figure 4.18.: Comparison of histograms of the BDT score for the background and signal testing and training subsamples, for the 'soft' BDT trained on the IceCube-dominated subsample after the second level of cuts. Both plots contain the same data, but the left plot has a linear y-scale and the right plot a logarithmic y-scale. KS p denotes the p-value that is the result of the Kolmogorov-Smirnov test described in section 4.3.3.

Sample	Rate after cut [mHz]	Passing fraction of cut
Data	0.373	0.291 %
Atmospheric μ simulation	0.039	0.026%
Atmospheric ν_μ simulation	0.37	14.5%
100 GeV bb signal simulation		37.5%
500 GeV bb signal simulation		54.3%

Table 4.6.: The rates and passing fractions for events which have a BDT score larger than 0.35 for the final level BDT for the DeepCore-dominated sample, for test data, atmospheric muon simulation, atmospheric muon neutrino simulation and two signal samples, where bb refers to the annihilation channel and the mass refers to the mass of the WIMP.

For both the hard and the soft BDT, there is reasonable agreement between data and simulation and there are no signs of either data-MC overtraining or training sample overtraining (as can be seen from the comparisons in figures 4.16 and 4.18). From this it can be concluded that both BDTs perform well and exhibit no problems.

After the final cut, the DeepCore-dominated and IceCube-dominated sample will be combined (which is possible because they are non-overlapping) and the resulting sample will be studied.

4.5. Properties of the final sample

The goal of the event selection described in the previous section is to remove as much of the atmospheric muon background as possible, while keeping as much of the well-reconstructed signal as possible. To quantify how well this goal has been reached, there are several variables that can be studied: the purity of the final-level sample (the fraction that consists of neutrinos, as opposed to atmospheric muons), the fraction of well-reconstructed signal events that is kept for this sample, and the median resolution of the direction reconstruction. As the event selection was designed for and tested only on IC86-2011 data, a comparison of the three years used in the event selection (IC86-2011, IC86-2012 and IC86-2013) is also shown.

4.5.1. Purity and signal efficiency

The cut on the final-level BDT score will be optimized for the best sensitivity. This is done separately for each of the targets (each of the dwarf galaxies, the combined dwarf galaxies, the M31 galaxy and the Virgo cluster) and will be discussed in more detail in chapter 5. For now, the effect of the different optimized BDT cuts is shown.

For the DeepCore-dominated sample, the optimal cut on the final-level BDT is the same for each target: 0.35. This value is denoted by the dashed green line in figure 4.13. The rates after this cut are given in table 4.6. Based on simulation and assuming that all neutrinos are atmospheric, the purity of the final sample is approximately 90%, meaning that there is still an atmospheric muon contamination of approximately 10%. As can be seen from figure 4.13, it is possible to remove more of the atmospheric muon background. However, this will lower the efficiency for signal neutrinos significantly, and for the analysis a small remaining atmospheric muon background is not a problem.

Sample	Pass fraction (soft, 0.275)	Pass fraction (soft, 0.3)
Data	1.37% (1.37 mHz)	0.783% (0.783 mHz)
Atmospheric μ simulation	0.20% (0.16 mHz)	0.11% (0.088 mHz)
Atmospheric ν_μ simulation	20.8% (1.02 mHz)	12.6% (0.62 mHz)
200 GeV $\mu\mu$ signal simulation	39.5%	26.2%
100 TeV $\mu\mu$ signal simulation	16.9%	10.7%

Table 4.7.: The rates and passing fractions for events which have a BDT score larger than 0.275 or larger than 0.3 for the final level soft BDT for the IceCube-dominated sample, for test data, atmospheric muon simulation, atmospheric muon neutrino simulation and two signal samples, where $\mu\mu$ refers to the annihilation channel and the mass refers to the mass of the WIMP.

Sample	Pass fraction (hard, 0.175)	Pass fraction (hard, 0.225)
Data	2.25% (2.25 mHz)	1.32% (1.32 mHz)
Atmospheric μ simulation	0.55% (0.44 mHz)	0.30% (0.24 mHz)
Atmospheric ν_μ simulation	30.8% (1.51 mHz)	18.5% (0.91 mHz)
200 GeV $\mu\mu$ signal simulation	16.2%	5.95%
100 TeV $\mu\mu$ signal simulation	83.5%	73.6%

Table 4.8.: The rates and passing fractions for events which have a BDT score larger than 0.175 or larger than 0.225 for the final level hard BDT for the IceCube-dominated sample, for test data, atmospheric muon simulation, atmospheric muon neutrino simulation and two signal samples, where $\mu\mu$ refers to the annihilation channel and the mass refers to the mass of the WIMP.

For the IceCube-dominated sample, there are two final-level BDTs. In both cases the optimization described in chapter 5 results in two possibilities for the final cut (depending on the target), so there are four possible final level samples. The final cut is placed at either 0.175 or 0.225 for the hard BDT and at either 0.275 or 0.3 for the soft BDT. These values are denoted by the dashed green lines in figures 4.15 and 4.17. The resulting rates and efficiencies are given in tables 4.7 and 4.8. Comparing these, one can see that the hard BDT performs better for the high-mass signal sample than the soft BDT, while for low-mass signal sample it is the other way around. Based on simulation, the purity of the final-level sample ranges from 77% to 88%.

From now on, the IceCube-dominated final level sample selected by a cut on the soft BDT (the BDT trained on a signal sample with relatively low neutrino energies) will be referred to as the 'soft IceCube-dominated sample' and the IceCube-dominated final level sample selected by a cut on the hard BDT (the BDT trained on signal samples with relatively high neutrino energies) will be referred to as the 'hard IceCube-dominated sample'.

4.5.2. Median angular resolution

Since the result of the analysis described in chapter 5 mainly depends on the quality of the reconstructed direction, the event selection has been designed to select relatively well-reconstructed events. Because this is only relevant for signal, in this section the resolution for signal events is shown. For 38 signal samples (19 WIMP masses and two annihilation channels: muons and b quarks) the median angular resolution at the final level of the event selection is calculated. These are plotted in figure 4.19, where a line is added to guide the eye, but the resolutions are only calculated for the masses denoted by the markers. As mentioned

before, there are several possible cuts for the IceCube-dominated sample. In this section, only events for which the soft BDT score is larger than 0.275 or the hard BDT score larger than 0.225 are considered, as in those cases the data rate in the final sample is approximately equal.

From figure 4.19, one can see that the resolution is better for higher WIMP masses. This is because SplineMPE performs better for high-energy events than for low-energy events (as was shown in figure 4.3), and signal samples for high WIMP masses contain more high-energy events than signal samples for low WIMP masses. Similarly, the resolution is better for annihilation to muons than for annihilation to b quarks, again because the former sample contains more high-energetic neutrinos (as was shown in figure 2.15). The resolution is also better for the IceCube-dominated sample than for the DeepCore-dominated sample, and better for the hard IceCube-dominated sample than for the soft IceCube-dominated sample. Again, this is mainly due to the different energy distributions of those samples, and to the fact that the hard BDT cuts more strongly on how well-reconstructed an event is.

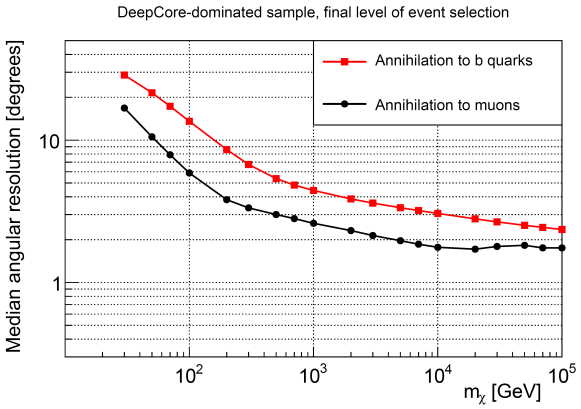
In figure 4.20, the median angular resolutions for annihilation to muons at the final level and at the penultimate level (before the final-level BDT cut) are shown. The fact that the errors are in general smaller at final level than at the penultimate level shows that the event selection indeed selects well-reconstructed events. The only exception is if the median resolution of the soft IceCube-dominated sample at masses above 10 TeV is compared to the sample before the final-level BDT, but at these WIMP masses the soft IceCube-dominated sample is not used (as will be shown in chapter 5). The soft BDT was not trained on these high-energy events, and is selecting relatively well-reconstructed events at lower energies (which can be similar to relatively badly-reconstructed events at higher energies).

4.5.3. Comparison of seasons

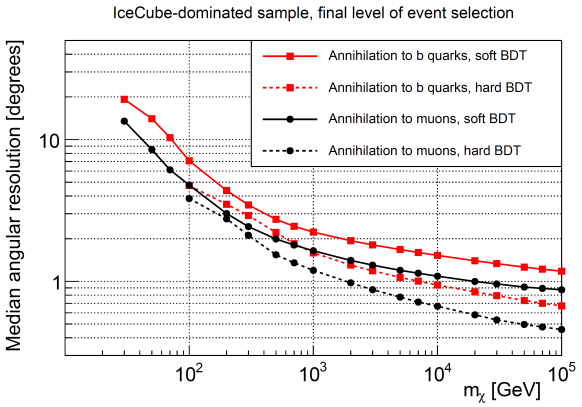
For this analysis, three years of IceCube data have been used, which are denoted by IC86-2011, IC86-2012 and IC86-2013. The event selection described in section 4.4 was designed using a subset of the IC86-2011 data and IC86-2011 simulation. To test whether it also performs well on IC86-2012 and IC86-2013 data, the same event selection is applied to these data sets. The resulting rates are shown in table 4.9. Comparisons of the distributions of the zenith as reconstructed by SplineMPE and the reconstructed energy (or energy proxy) are shown in figures 4.21 (DeepCore-dominated sample, events with a BDT score larger than 0.35), 4.22 (IceCube-dominated sample, events with a soft BDT score larger than 0.275) and 4.23 (IceCube-dominated sample, events with a hard BDT score larger than 0.175).

From this one can see that there is a difference in rate between IC86-2011 on one side and IC86-2012 and IC86-2013 on the other side, while there are no statistically significant differences between IC86-2012 and IC86-2013. This is likely related to the changes in calibration and filtering between IC86-2011 and IC86-2012 that were discussed in sections 4.4.1 and 4.4.2. However, the difference between IC86-2011 and the other two years is small, and it was decided that the possible small improvement to the sensitivity does not outweigh the work involved in redoing the event selection for IC86-2012 and IC86-2013. This means that the same event selection was used for all three years.

To test whether the differences between IC86-2011 and IC86-2012 also show up in the signal simulation, the event selection was also applied to an IC86-2012 simulated neutrino dataset.

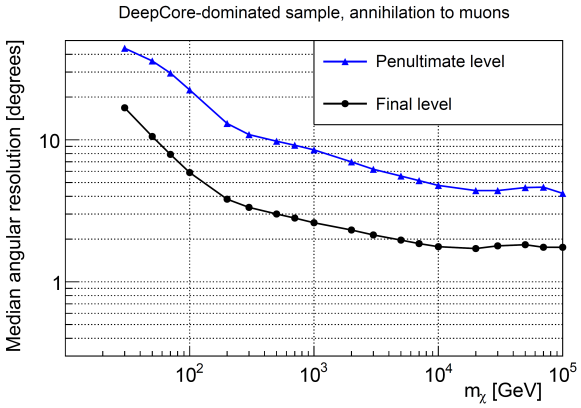


(a) DeepCore-dominated, events with a BDT score higher than 0.35

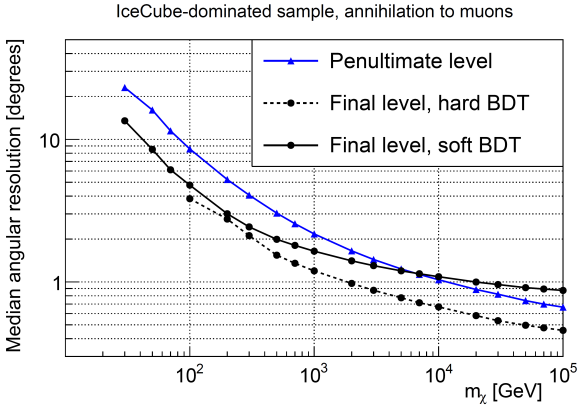


(b) IceCube-dominated, events with a BDT score higher than 0.275 (for the soft BDT) or 0.225 (for the hard BDT)

Figure 4.19.: The median angular resolution for different WIMP masses and annihilation channels, using simulated signal events passing the event selection up to the final level. The lines are added to guide the eye, but only the masses denoted by the markers (dots or squares) are tested.



(a) DeepCore-dominated, where 'final level' refers to events with a BDT score higher than 0.35

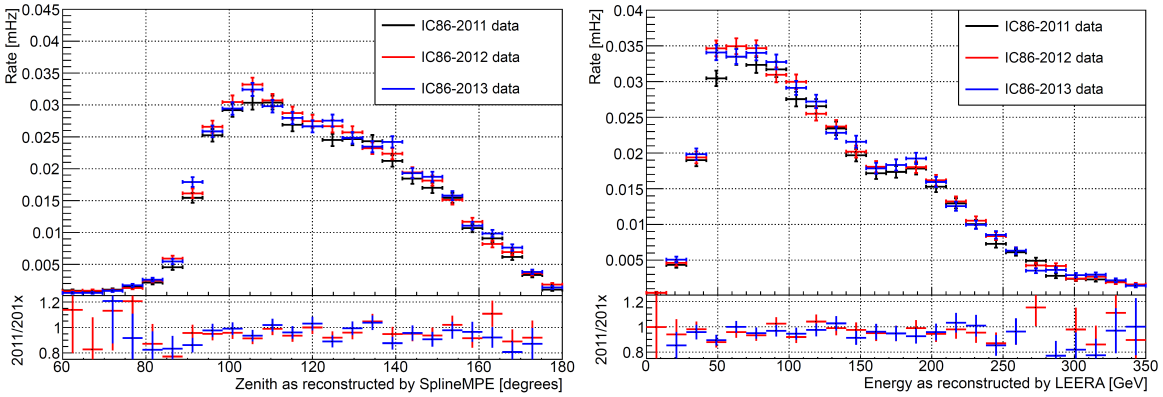


(b) IceCube-dominated, where 'final level' refers to events with a BDT score higher than 0.275 (for the soft BDT) or 0.225 (for the hard BDT)

Figure 4.20.: The median angular resolution for WIMPs of different masses annihilating to muons, using either simulated signal events passing the event selection up to the final level, or up to the penultimate level. The lines are added to guide the eye, but only the masses denoted by the markers (dots or triangles) are tested.

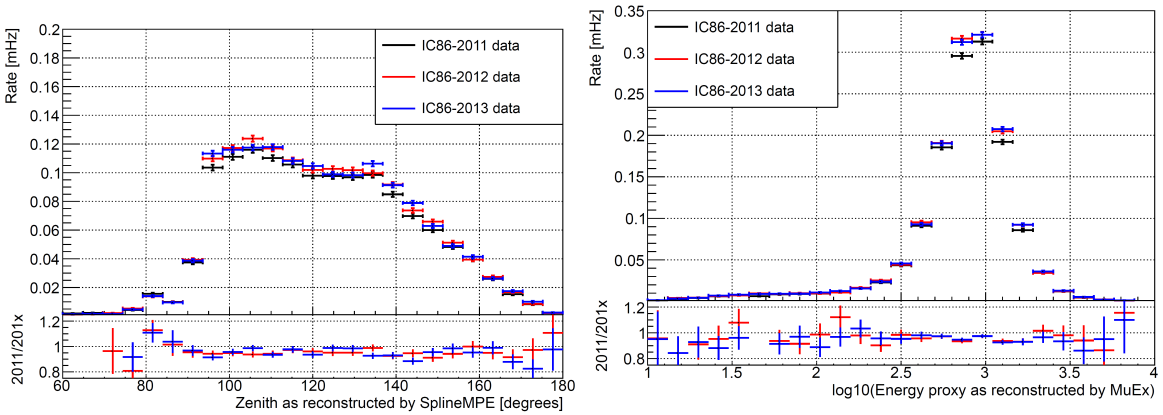
Sample	IC86-2011 data rate	IC86-2012 data rate	IC86-2013 data rate
DC, score $>$ 0.35	0.373 $\pm$ 0.004 mHz	0.390 $\pm$ 0.004 mHz	0.391 $\pm$ 0.004 mHz
IC, soft score $>$ 0.275	1.369 $\pm$ 0.007 mHz	1.436 $\pm$ 0.007 mHz	1.437 $\pm$ 0.007 mHz
IC, soft score $>$ 0.3	0.783 $\pm$ 0.005 mHz	0.817 $\pm$ 0.005 mHz	0.832 $\pm$ 0.005 mHz
IC, hard score $>$ 0.175	2.251 $\pm$ 0.009 mHz	2.332 $\pm$ 0.009 mHz	2.336 $\pm$ 0.009 mHz
IC., hard score $>$ 0.225	1.321 $\pm$ 0.007 mHz	1.355 $\pm$ 0.007 mHz	1.365 $\pm$ 0.007 mHz

Table 4.9.: The data rates for events at the final level of the event selection, for the three years of data used in the analysis (IC86-2011, IC86-2012, IC86-2013), for all cuts on the output score of the final BDT that are used. DC refers to the DeepCore-dominated sample, IC refers to the IceCube-dominated sample.



(a) Zenith as reconstructed by SplineMPE **(b)** Energy as reconstructed by LEERA

Figure 4.21.: Comparing the distributions of zenith as reconstructed by SplineMPE and energy as reconstructed by LEERA for IC86-2011, IC86-2012 and IC86-2013, for data events in the DeepCore-dominated sample which have a final BDT score larger than 0.35.



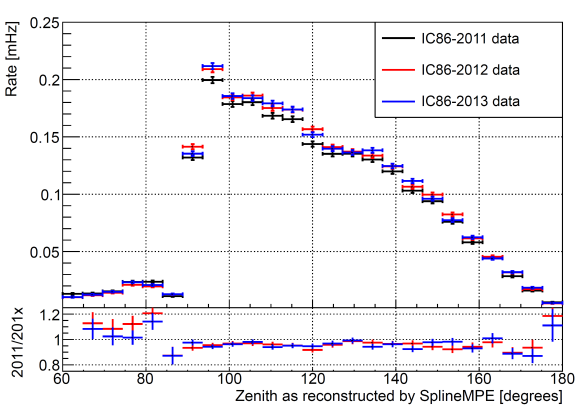
(a) Zenith as reconstructed by SplineMPE **(b)** The logarithm of the energy proxy as reconstructed by MuEx

Figure 4.22.: Comparing the distributions of zenith as reconstructed by SplineMPE and the logarithm of the energy proxy as reconstructed by MuEx for IC86-2011, IC86-2012 and IC86-2013, for data events in the IceCube-dominated sample which have a final soft BDT score larger than 0.275.

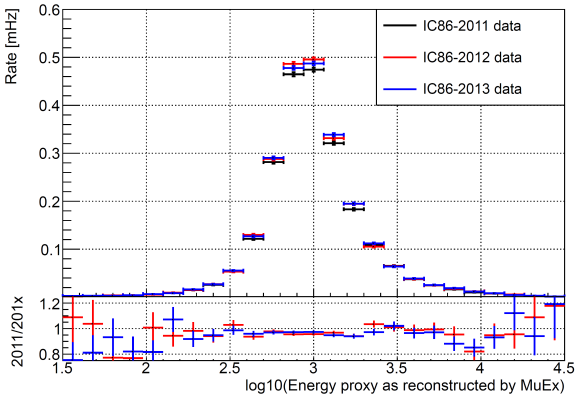
Sample	IC86-2011 total weight	IC86-2012 total weight
DC, BDT score > 0.35	2.35 ± 0.04	2.30 ± 0.04
IC, soft BDT score > 0.275	57.9 ± 0.2	57.7 ± 0.2
IC, soft BDT score > 0.3	33.4 ± 0.2	33.3 ± 0.2
IC, hard BDT score > 0.175	264.7 ± 0.5	264.2 ± 0.5
IC, hard BDT score > 0.225	216.3 ± 0.4	215.7 ± 0.4

Table 4.10.: The total weight (weighting each event with E_ν^{-2} and assuming an isotropic flux) for simulated neutrino events at the final level of the event selection, comparing IC86-2011 and IC86-2012, for all cuts on the output score of the final BDT which are used. DC refers to the DeepCore-dominated sample, IC refers to the IceCube-dominated sample.

For the comparison, an isotropic flux was assumed and each event was weighted with E_ν^{-2}. The total weight of all events passing the event selection (in arbitrary units) is compared in table 4.10. Comparisons of the distributions of the zenith as reconstructed by SplineMPE and the reconstructed energy (or reconstructed energy proxy) are shown in figures 4.24 (DeepCore-dominated sample, events with a BDT score larger than 0.35), 4.25 (IceCube-dominated sample, events with a soft BDT score larger than 0.275) and 4.26 (IceCube-dominated sample, events with a hard BDT score larger than 0.175). There is no significant difference between IC86-2011 and IC86-2012, from which it can be concluded that the difference in the data rates must be due to atmospheric muons, not to neutrinos. It also means that it is valid to use the same signal simulation (produced assuming the settings for the IC86-2011 season) for all three seasons.

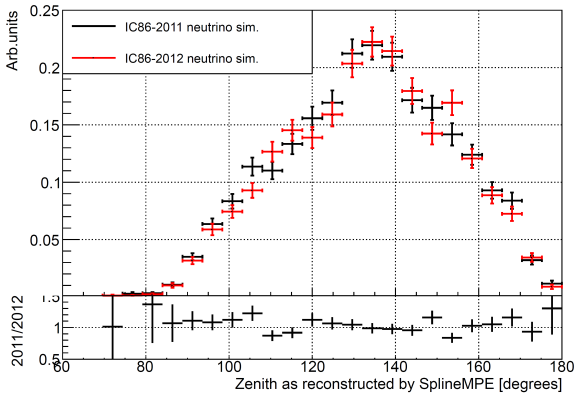


(a) Zenith as reconstructed by SplineMPE

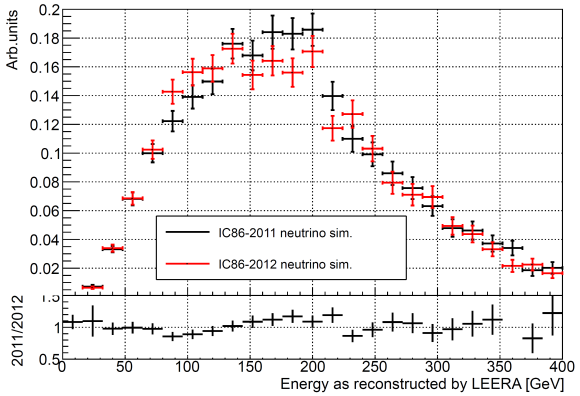


(b) The logarithm of the energy proxy as reconstructed by MuEx

Figure 4.23.: Comparing the distributions of zenith as reconstructed by SplineMPE and the logarithm of the energy proxy as reconstructed by MuEx for IC86-2011, IC86-2012 and IC86-2013, for data events in the IceCube-dominated sample which have a final hard BDT score larger than 0.175.

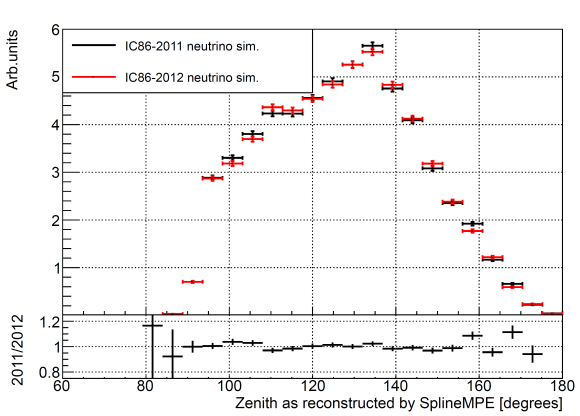


(a) Zenith as reconstructed by SplineMPE

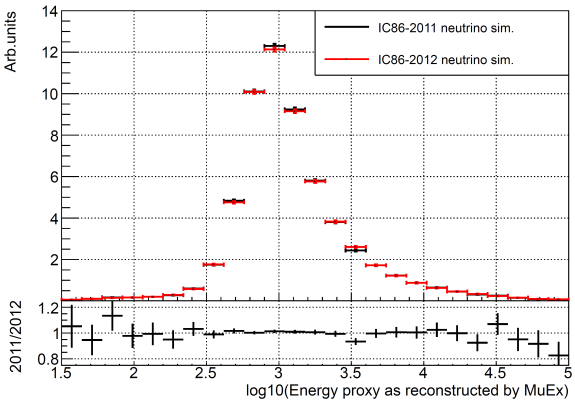


(b) Energy as reconstructed by LEERA

Figure 4.24.: Comparing the distributions of zenith as reconstructed by SplineMPE and energy as reconstructed by LEERA for IC86-2011 and IC86-2012 neutrino simulation (assuming an isotropic flux and weighting the events with E_ν^{-2}), for the events in the DeepCore-dominated sample which have a final BDT score larger than 0.35.

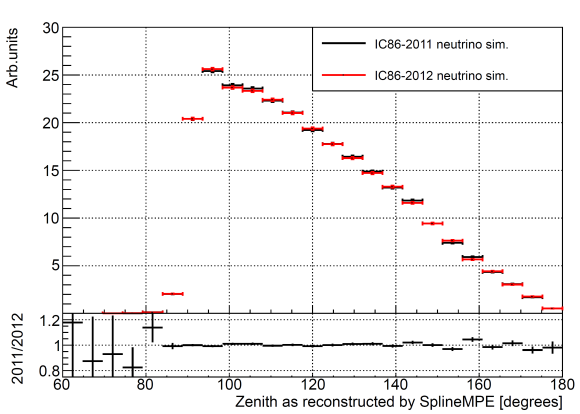


(a) Zenith as reconstructed by SplineMPE

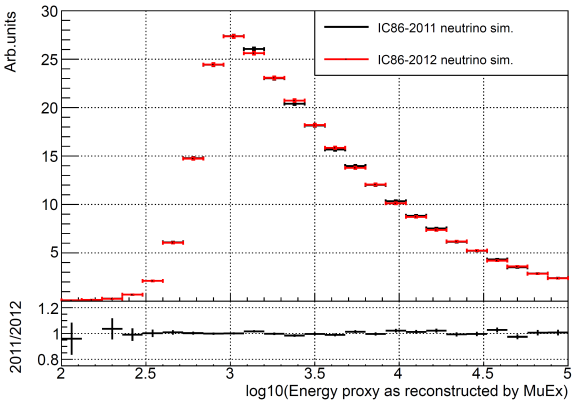


(b) The logarithm of the energy proxy as reconstructed by MuEx

Figure 4.25.: Comparing the distributions of zenith as reconstructed by SplineMPE and logarithm of the energy proxy as reconstructed by MuEx for IC86-2011 and IC86-2012 neutrino simulation (assuming an isotropic flux and weighting the events with E_ν^{-2}), for the events in the IceCube-dominated sample which have a final soft BDT score larger than 0.275.



(a) Zenith as reconstructed by SplineMPE



(b) The logarithm of the energy proxy as reconstructed by MuEx

Figure 4.26.: Comparing the distributions of zenith as reconstructed by SplineMPE and logarithm of the energy proxy as reconstructed by MuEx for IC86-2011 and IC86-2012 neutrino simulation (assuming an isotropic flux and weighting the events with E_ν^{-2}), for the events in the IceCube-dominated sample which have a final hard BDT score larger than 0.175.

5. Analysis method

The event selection described in the previous chapter results in a sample that is dominated by neutrinos, with a small contamination of atmospheric muons (10-20%, depending on which BDT cut is used). These neutrinos are mainly atmospheric in origin, but some of them might originate outside the atmosphere of the Earth, for example in dark matter annihilations in nearby galaxies and galaxy clusters. The goal of the analysis method described in this chapter is to determine whether or not this is the case, and if not, what upper limit can be placed on the dark matter annihilation cross section. This is done by comparing the observed data to what one would expect if there were no signal. In that case, only background is left, which is known to be approximately isotropic and has a known energy distribution which is different from that for signal events.

One could do this using a binned analysis, where a signal region is defined, for example a circular region in space around the target. All events inside this region are counted and compared to the background expectation (which is estimated from a control region), and if there is an excess, its significance is calculated using for example Poissonian statistics. This method was used in the previous IceCube analysis which looked for a dark matter signal from nearby galaxies and galaxy clusters [133]. For the analysis described in this thesis, a more sophisticated method was used: an unbinned maximum likelihood method. Using this, one can give a higher weight to events that are more signal-like and a lower weight to events that are less signal-like. This type of method has been shown to perform significantly better for similar IceCube analyses [211], and it was improved further specifically for this analysis by also including the PDF for the estimated error of the direction reconstruction.

The unbinned maximum likelihood method will be described in section 5.1. Some improvements to the method, allowing it to perform better for this analysis, are described in sections 5.2 and 5.3. In section 5.4, the adaptations to the method for extended sources (instead of just point sources) are detailed. Finally, in section 5.5, the method to determine the sensitivity to the velocity-averaged WIMP annihilation cross section is described, and in section 5.6, the resulting sensitivities for the analysis described in this thesis are shown.

5.1. Unbinned maximum likelihood method

The basic idea of an unbinned maximum likelihood method can be understood using the following example: consider a parameter θ which has different distributions for signal and background. These distributions are normalized and denoted by $S(\theta)$ and $B(\theta)$, and are usually called Probability Density Functions or PDFs. Now assume that the sample used for the analysis consists of n_{obs} events in total, of which μ are signal events. In that case, the probability to get an event with $\theta = \theta_i$ is

$$f(\theta_i|\mu) = \frac{\mu}{n_{obs}} \cdot S(\theta_i) + \left(1 - \frac{\mu}{n_{obs}}\right) \cdot B(\theta_i). \tag{5.1}$$

Since the events in the sample are independent of each other, the probability (the likelihood) to get a sample with a set of values $\{\theta_i\}$ is the product of the probabilities for the single events:

$$\begin{aligned}\mathcal{L}(\{\theta_i\}|\mu) &= \prod_{i=1}^{n_{obs}} f(\theta_i|\mu) \\ &= \prod_{i=1}^{n_{obs}} \frac{\mu}{n_{obs}} \cdot S(\theta_i) + \left(1 - \frac{\mu}{n_{obs}}\right) \cdot B(\theta_i). \end{aligned} \tag{5.2}$$

By maximizing $\mathcal{L}(\{\theta_i\}|\mu)$ as a function of μ, one obtains the best fit for the number of signal events, which does not necessarily have to result in an integer number. In the implementation used for this analysis (where the logarithm of $\mathcal{L}$ is used for calculations), μ is required to be non-negative and smaller than n_{obs}. This is acceptable since a negative number of signal events or a number of signal events larger than the total number of events does not have a physical interpretation.

Equation 5.2 can straightforwardly be extended to multiple parameters, for example to two parameters: θ and ϕ. If these parameters are independent, the probability to get values θ_i and ϕ_i for a signal event is $S(\theta_i, \phi_i) = f_S(\theta_i) \cdot g_S(\phi_i)$, and likewise for B. If they are not independent, two-dimensional PDFs $S(\theta_i, \phi_i)$ and $B(\theta_i, \phi_i)$ are used.

It is also possible to extend the method to multiple independent samples, for example the IceCube-dominated and the DeepCore-dominated subsamples and the different years of data-taking (IC86-2011, IC86-2012 and IC86-2013), which can all have different distributions for the parameters. This is done by multiplying the likelihoods, for j independent samples and one parameter θ this results in

$$\begin{aligned}\mathcal{L}(\{\theta_i\}|\mu) &= \prod_j \prod_{i=1}^{n_{obs,j}} f_j(\theta_i|\mu_j) \\ &= \prod_j \prod_{i=1}^{n_{obs,j}} \frac{\mu_j}{n_{obs,j}} \cdot S_j(\theta_i) + \left(1 - \frac{\mu_j}{n_{obs,j}}\right) \cdot B_j(\theta_i) \end{aligned} \tag{5.3}$$

with the number of signal events in the jth sample determined by

$$\mu_j = \frac{w_j}{w}\mu \tag{5.4}$$

where w is the sum of the signal weights of all signal events in all samples combined and w_j is the sum of the weights of all signal events in the jth sample. As a reminder, each simulated signal event is given a signal weight in such a way that the resulting weighted energy spectrum and rate match the energy spectrum and rate of the signal sample that is being simulated (for example, 100 GeV WIMPS annihilating to muons in the Virgo cluster), see section 3.4.

The final result of the maximum likelihood method is a fitted number of signal events in the data sample. Since the likelihood is calculated using distributions of parameters which depend on the specific signal sample considered, the sample must be tested for each signal

sample separately. The trial factor resulting from many separate tests will be discussed in section 6.1.

5.1.1. Probability density functions

The likelihood given in equation 5.2 can be calculated and optimized once the probability density functions S and B are known. In the case of the analysis described in this thesis, three variables are used:

- the reconstructed energy or energy proxy E, as reconstructed by LEERA for the DeepCore-dominated sample, or as reconstructed by MuEx for the IceCube-dominated sample (both reconstructions were discussed in section 4.1.4),
- the angle Ψ between the reconstructed direction as reconstructed by SplineMPE (which was discussed in section 4.1.2) and the direction of the target (Ψ is often called the space angle),
- the estimated error on the direction reconstruction σ as reconstructed by Paraboloid (which was discussed in section 4.1.3).

Traditionally, in IceCube analyses only the space angle and the reconstructed energy or energy proxy are used [211]. However, as will be shown below, the PDF for the space angle Ψ depends on the estimated error σ. This means that if the distributions for σ are different for signal and background (which is the case for the analysis described in this thesis), the PDF for σ needs to be included in the likelihood [212].

When doing an unbinned maximum likelihood analysis, in general the product in equation 5.2 runs over all events in the experimental data sample. However, in the case discussed here, events which are far from the target will contribute little, since the probability to get a signal event with a large space angle is small. That means that removing these events from the sample will have a negligible effect on the final result, but it will significantly speed up the calculations. As the background PDFs will be based on data events which are given a random reconstructed azimuth, events with an azimuth far from the target should be kept, while events with a zenith far from the target can be removed.

In practice, what is done is the following: for each of the five dwarf galaxies, only events where the difference between the reconstructed zenith and the zenith of the target is less than 2.5 degrees are considered. For the M31 galaxy and the Virgo cluster, due to the possibility of an extended signal flux, a wider zenith band is used: events where the difference between the reconstructed zenith and the zenith of the target is less than 5 degrees are considered.

An additional advantage of only considering events in a 5-degree-wide zenith band centered on each of the dwarf galaxies is that these bands do not overlap for the five dwarf galaxies considered in this analysis. This means that these samples are independent, so they can be combined using the method of equation 5.3.

Background

As has been mentioned before, experimental data is used as the background sample for the final analysis, instead of Monte Carlo. To avoid a bias (since the data could contain signal

events), the data is scrambled before it is used as background: each data event is assigned a random azimuth. This is possible because there is no significant dependence on azimuth for background events, due to the (approximate) symmetry of the detector. If there is a signal component in the data, it will be washed out by the scrambling, so the scrambled data can safely be used as a background sample.

For background events there is no correlation between the space angle Ψ and the other two variables. However, there is a correlation between reconstructed energy E and estimated error σ, since both are correlated to the true neutrino energy. This means that the background PDF can be written as

$$B(\Psi, E, \sigma) = f_B(\Psi) \cdot g_B(E, \sigma). \tag{5.5}$$

It is possible that the distributions for E, Ψ or σ depend on the zenith. For example, the acceptance of the event selection is zenith-dependent, and the background rate is also larger close to the horizon than further away. To take this into account, both $f_B(\Psi)$ and $g_B(E, \sigma)$ are determined separately for each target (the five dwarf galaxies, the M31 galaxy and the Virgo cluster). In each case, all events inside the zenith band of this target are used to determine the PDF.

For $f_B(\Psi)$, the PDF is smoothed by using each data event 1000 times, each time with a different (random) azimuth so the calculated Ψ is different as well. The effect of this smoothing on the PDF for the Segue dwarf galaxy for the DeepCore-dominated sample at final level is shown in figure 5.1. In this figure one can also see the distinct shape of $f_B(\Psi)$, which is a consequence of only using events in the zenith band around the target: since the background is approximately isotropic, $f_B(\Psi)$ is proportional to the size of the overlap between the zenith band and a ring of infinitesimal width at distance Ψ from the target. For $g_B(E, \sigma)$, the two-dimensional histogram is smoothed by using Kernel Density Estimation, which is described in section 5.3.

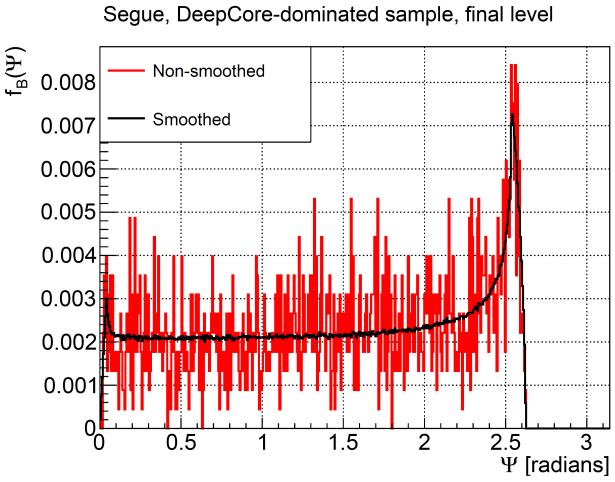


Figure 5.1.: Comparing $f_B(\Psi)$ for the non-smoothed (using each event once) and the smoothed (using each event 1000 times, each time with a different random azimuth) case, for the Segue dwarf galaxy and DeepCore-dominated sample at final level.

Signal

As has been discussed in section 3.4, the simulated signal sample is not a true point source sample (where all neutrinos come from the direction of the target that is being considered) but contains neutrinos from all directions. To get the signal sample that is used for the PDFs, two cuts are applied:

- All events where the difference between the true zenith (as simulated) and the zenith of the target that is being considered is larger than 5 degrees are removed. This is done because distributions of event properties can be quite different for different zenith ranges, and the goal is to have a sample that is as similar as possible to a true point source sample.

- All events where the difference between the true zenith and the reconstructed zenith is larger than 2.5 degrees are removed. If these signal events were true signal events coming from the direction of the target that is being considered, they would not be taken into account in the analysis, so they should not contribute to the PDF either.

For the signal space angle distribution $f_S(\Psi)$, one could use the distribution of angular differences between the true direction and the reconstructed direction. That would mean that each event is treated equally: for each event, the same PDF is used. However, some events are known to be better reconstructed than others (as they have a lower estimated error of the direction reconstruction σ). It would improve the final result if these events could have a larger influence than events with a higher σ. To take this into account, a normalized two-dimensional Gaussian with a width equal to σ is used as PDF:

$$f_S(\Psi) = \frac{\Psi}{\sigma^2} e^{-\frac{\Psi^2}{2\sigma^2}}. \tag{5.6}$$

To get the full PDF $S(\Psi, E, \sigma)$ for signal, equation 5.6 needs to be multiplied with the two-dimensional PDF for reconstructed energy E and estimated error σ, $g_S(E, \sigma)$. This is the two-dimensional histogram of the simulation sample described above, smoothed using Kernel Density Estimation, which is described in section 5.3. The final PDF for signal is then given by

$$S(\Psi, E, \sigma) = \frac{\Psi}{\sigma^2} e^{-\frac{\Psi^2}{2\sigma^2}} \cdot g_S(E, \sigma). \tag{5.7}$$

5.2. Recalibration of paraboloid sigma

The Paraboloid sigma σ needs to be recalibrated before it is used to determine the signal PDF, as the result of the Paraboloid algorithm generally underestimates the true error (as was shown in figure 4.4). This recalibration is done by multiplying σ with a recalibration factor which depends on the number of hit DOMs (for the DeepCore-dominated sample) or the logarithm of the energy proxy as reconstructed by MuEx (for the IceCube-dominated sample).

To determine how the Paraboloid sigma should be recalibrated, the event sample at the penultimate level (before the final BDT cut) is used, to have enough statistics. As the true energy spectrum could influence the results of this method, the events are weighted with a mixture of several signal samples (the sum of all weights for a signal sample is normalized to one, and the sum of the normalized weights for each event is used), all for annihilation to

muons, but with different WIMP masses. To determine the recalibration for the DeepCore-dominated sample, signal samples for WIMP masses of 30, 50, 70 and 100 GeV are used. For the soft IceCube-dominated sample, signal samples for WIMP masses of 100, 300, 500 and 700 GeV are used. For the hard IceCube-dominated sample, signal samples for WIMP masses of 1 TeV, 10 TeV and 100 TeV are used. As a reminder, the 'soft' or 'hard' IceCube-dominated sample refers to the final-level IceCube-dominated sample resulting from a cut on the output of the soft or hard BDT.

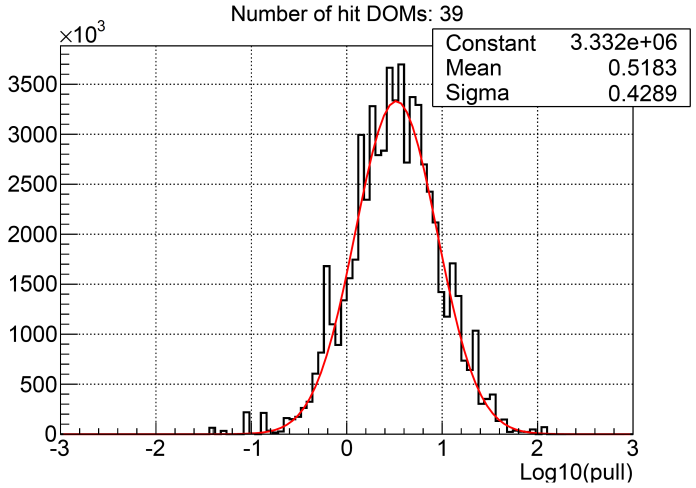


Figure 5.2.: The distribution of log10(pull) values, where the pull is defined as the true error divided by the Paraboloid sigma, for all events in the simulated DeepCore-dominated sample which have 39 hit DOMs. This distribution is fitted with a Gaussian, and the fitted mean of this Gaussian is used for further calculations.

The following method is used to determine the formula for recalibration of the Paraboloid sigma: the events are divided in bins of the energy-related variable (with width 1 for the number of hit DOMs and 0.2 for the logarithm of the energy proxy as reconstructed by MuEx). For each bin, a histogram of the logarithm of the pull values of all events in this bin is plotted and fitted with a Gaussian (this is shown for one bin in figure 5.2). The pull is defined as the true error divided by the Paraboloid sigma. The 'average pull' is then defined as

$$p_{av} = 10^{\mu} \tag{5.8}$$

where μ is the center of the Gaussian (as resulting from the fit). When defined like this, p_{av} is similar to the mean pull in the bin, but outliers have less influence on this average pull than on the mean pull. Then, p_{av} is plotted as a function of the energy-related variable. For the range of values of the energy-related variable where there are at least ten events in a bin, this is fitted with a polynomial. This polynomial can now be used to recalibrate σ.

The resulting plots of the 'average' pull against the energy-related variable, including the polynomial fit (in blue), are shown in figures 5.3 (for the DeepCore-dominated sample), 5.4 (for the IceCube-dominated soft sample) and 5.5 (for the IceCube-dominated hard sample).

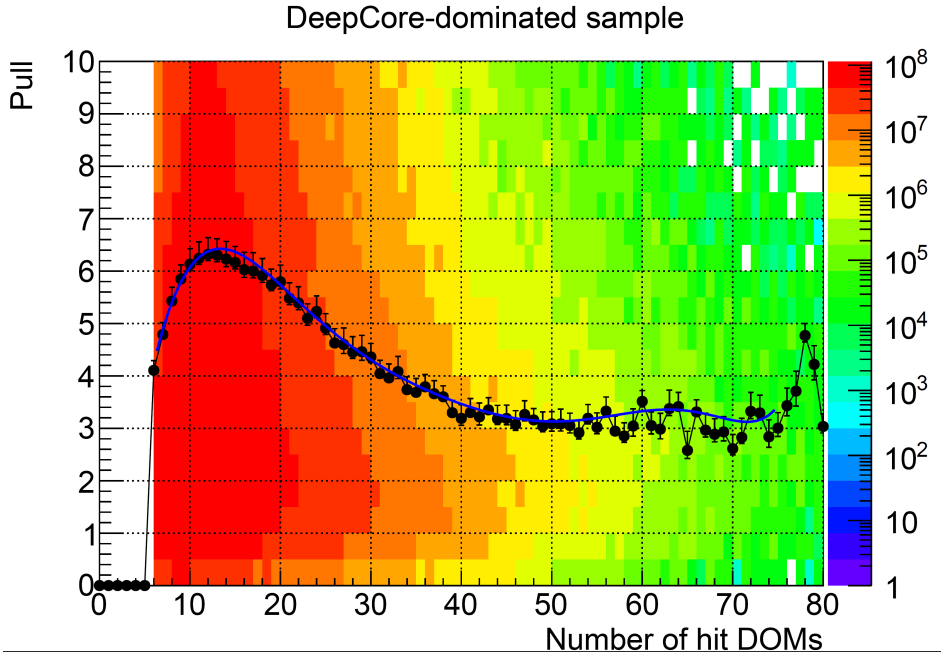


Figure 5.3.: This shows the two-dimensional histogram of the pull and the number of hit DOMs for the DeepCore-dominated sample, the average pull p_{av} as a function of the number of hit DOMs in black, and the polynomial fit to p_{av} in blue.

The polynomial fit to the average pull as a function of the energy-related variable can now be used to recalibrate the Paraboloid sigmas. For the DeepCore-dominated sample, this is

$$\sigma_{recal} = \sigma_{orig} \cdot \sum_{i=0}^{8} a_i N_{ch}^i \tag{5.9}$$

with σ_{recal} the recalibrated Paraboloid sigma, σ_{orig} the original Paraboloid sigma (the output of the algorithm described in section 4.1.3), N_{ch} the number of hit DOMs (after cleaning) and $a_i = \{3.713, 2.043, \text{-}0.1408, 3.491 \cdot 10^{-3}, 1.678 \cdot 10^{-5}, \text{-}2.813 \cdot 10^{-6}, 6.268 \cdot 10^{-8}, \text{-}6.035 \cdot 10^{-10}, 2.214 \cdot 10^{-12}\}$. Equation 5.9 is plotted in figure 5.3 in blue.

For the IceCube-dominated soft sample, the polynomial fit is

$$\sigma_{recal} = \sigma_{orig} \cdot \sum_{i=0}^{6} b_i \cdot (\log E)^i \tag{5.10}$$

with $\log E$ the logarithm of the energy proxy as reconstructed by MuEx and $b_i = \{\text{-}26.32, 67.08, \text{-}48.25, 11.0, 1.578, \text{-}1.008, 0.1119\}$. Equation 5.10 is plotted in figure 5.4 in blue.

Finally, for the IceCube-dominated hard sample, the polynomial fit is

$$\sigma_{recal} = \sigma_{orig} \cdot \sum_{i=0}^{6} c_i \cdot (\log E)^i \tag{5.11}$$

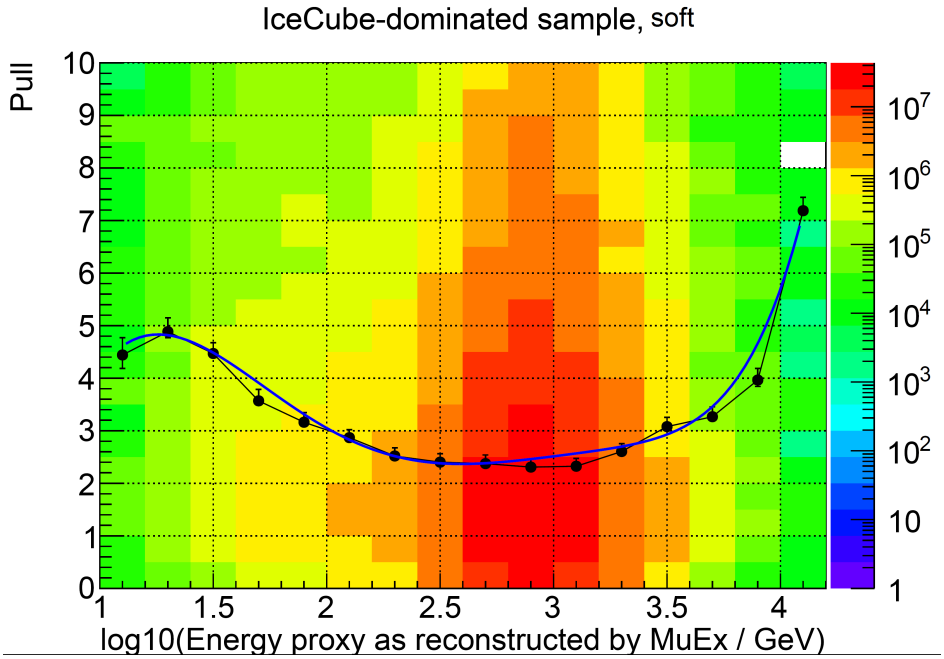


Figure 5.4.: This shows the two-dimensional histogram of the pull and the log10 of the energy proxy as reconstructed by MuEx for the IceCube-dominated soft sample, the average pull p_{av} as a function of the number of hit DOMs in black, and the polynomial fit to p_{av} in blue.

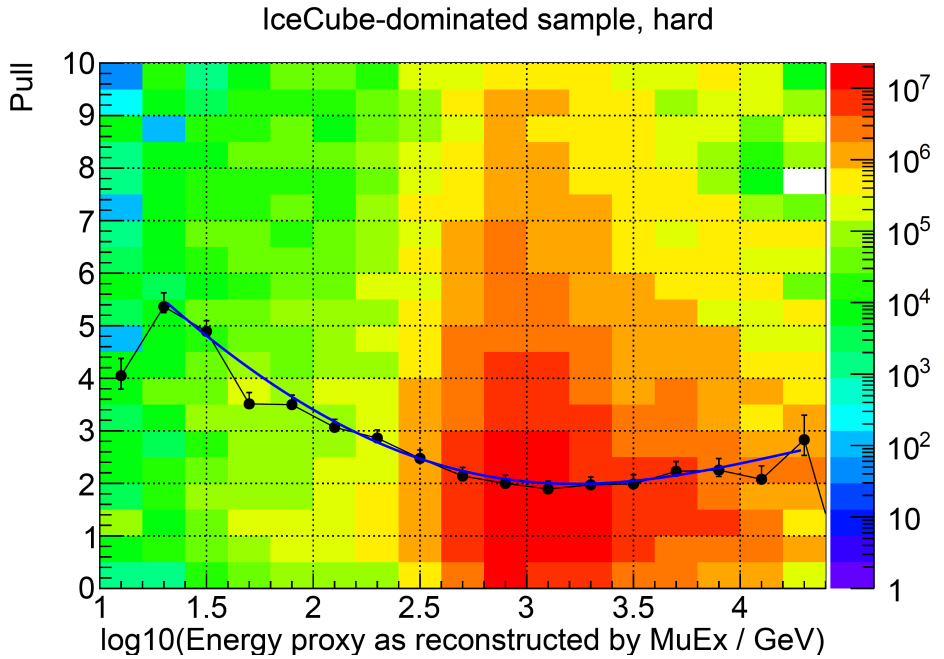


Figure 5.5.: This shows the two-dimensional histogram of the pull and the log10 of the energy proxy as reconstructed by MuEx for the IceCube-dominated hard sample, the average pull p_{av} as a function of the number of hit DOMs in black, and the polynomial fit to p_{av} in blue.

with $c_i = \{10.86, -4.382, -0.07311, 0.1949, 0.01573, -7.293 \cdot 10^{-3}, 2.877 \cdot 10^{-4}\}$. Equation 5.11 is plotted in figure 5.5 in blue.

5.3. Kernel Density Estimation

As has been discussed in section 5.1.1, a two-dimensional PDF is used for the energy proxy and the (recalibrated) Paraboloid sigma, both for signal and for background. Especially for the experimental data, the statistics are low. For example, in the DeepCore-dominated sample at final level (after removing all events with a final BDT score lower than 0.35) there are approximately 20,000 events left per year of data, but only a few hundred to a few thousand of these will be inside the zenith band for any specific target. This leads to significant statistical fluctuations and binning effects if a reasonable number of bins is chosen for a histogram (for example 10 bins in both energy proxy and Paraboloid sigma), an example of this is shown in figure 5.6a.

This problem can be solved by applying Kernel Density Estimation [213, 214]. Instead of using a histogram, the PDF is taken to be the sum of kernels centered at each data point (each event in the sample). In the analysis described in this thesis, these kernels are taken to be two-dimensional Gaussians, so the PDF is

$$g(\log E, \log \sigma) = \sum_{i=0}^{n_{obs}} w_i \cdot e^{-\frac{1}{2}\left(\frac{(\log E - \log E_i)^2}{b_E^2} + \frac{(\log \sigma - \log \sigma_i)^2}{b_\sigma^2}\right)} \tag{5.12}$$

with n_{obs} the number of events in the sample, w_i the weight of the ith event (in case of experimental data, all weights are 1), $\log E_i$ and $\log \sigma_i$ the logarithms of the energy proxy and recalibrated Paraboloid sigma for the ith event, and b_E and b_σ the bandwidths for both variables (the width of the Gaussians). These bandwidths do not have a physical meaning, but can be compared to for example the bin width in a histogram: they should be chosen in such a way that the resulting PDF has the required level of detail to show actual features as opposed to statistical fluctuations (see below for how the bandwidths are determined). The bandwidths for both variables (b_E and b_σ) are assumed not to be correlated. Equation 5.12 results in a smooth PDF at every value for the input variables.

Picking the best values for the bandwidths is the main challenge when using Kernel Density Estimation. For the analysis described in this thesis, the so-called 'Silverman's rule of thumb' was used [215], which states that the bandwidth for a variable j is

$$b_j = \left(\frac{n_{obs}(d+2)}{4}\right)^{-\frac{1}{d+4}} \sigma_j \tag{5.13}$$

where n_{obs} is the number of events in the sample, d is the dimensionality of the PDF (in this case, $d = 2$) and σ_j is the standard deviation of variable j.

Silverman's rule of thumb results in a bandwidth that is the same for each point. However, to get a PDF which represents the data as closely as possible, it would be preferable to have larger bandwidths in areas where there are few data points and smaller bandwidths in areas where there are many data points. To improve the result from Silverman's rule of thumb, an adaptive method is used [216]: the bandwidth is multiplied with a factor λ_i which is different

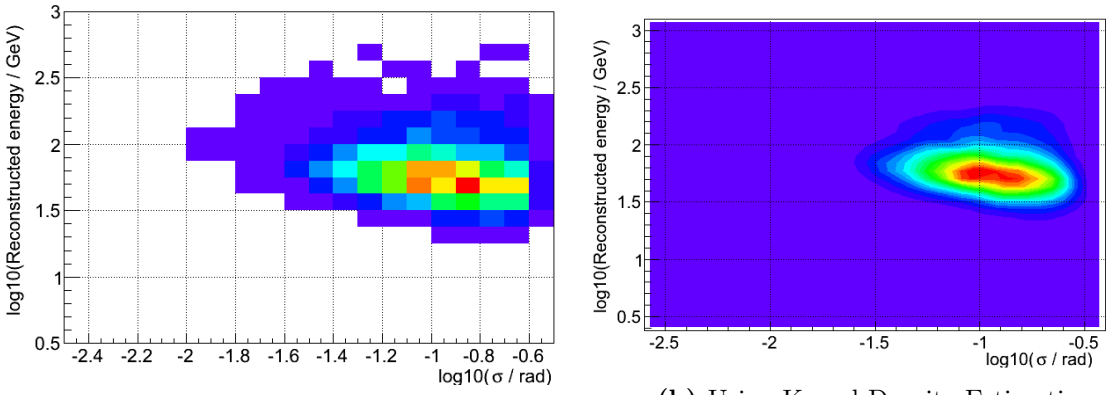
for each event. It is calculated as follows:

$$\lambda_i = \left(\frac{g(\log E_i, \log \sigma_i)}{A} \right)^{-0.3} \tag{5.14}$$

where $g(\log E_i, \log \sigma_i)$ is the value of the first iteration of the KDE (using the bandwidths given by equation 5.13) at the location of point i. A can be thought of as the 'average' KDE value for each of the data points and is given by

$$A = e^{\frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} \ln \left(g(\log E_i, \log \sigma_i) \right)} \tag{5.15}$$

where $g(\log E_i, \log \sigma_i)$ has been defined above. This means that if the first-iteration KDE value at a certain point is smaller than the 'average' KDE value (which happens when there are few events in that area, or only low-weight events), the λ_i value is larger, so in the second iteration a wider Gaussian is used for this data point. If on the other hand the first-iteration KDE value is larger than the 'average', the λ_i value will be smaller, so a narrower Gaussian will be used in the second iteration.



(a) Two-dimensional histogram

(b) Using Kernel Density Estimation

Figure 5.6.: The two-dimensional PDF for a simulated signal sample (100 GeV WIMPs annihilating to b quarks, DeepCore-dominated sample at final level), using a two-dimensional histogram (a) and using Kernel Density Estimation (b).

Applying the method described above on the same data that was used to create the two-dimensional histogram shown in figure 5.6a results in a smooth PDF, as shown in figure 5.6b. The KDE method is applied to all two-dimensional signal and background PDFs that are used in this analysis.

5.4. Extended sources

The method described in section 5.1 has been designed for point sources, where all signal events originate from the same point on the sky. However, as was discussed in section 2.5, two of the targets considered here (the M31 galaxy and the Virgo cluster) could have a more extended emission due to dark matter substructure. To test that case, the method is adapted slightly. The spatial signal PDF $f_S(\Psi)$ is no longer described by the two-dimensional Gaussian with width equal to the recalibrated Paraboloid sigma (equation 5.6), but by the convolution of this two-dimensional Gaussian with the spatial emission profile of the target.

In the case that is considered here, the spatial emission profile of the target is not known, as the dark matter profile including substructure is not known. A two-dimensional Gaussian profile is assumed, and the result of convoluting this profile with the two-dimensional Gaussian with width equal to the recalibrated Paraboloid sigma is a two-dimensional Gaussian with width

$$\sigma_w = \sqrt{\sigma_P^2 + \sigma_s^2}, \tag{5.16}$$

with σ_P the recalibrated Paraboloid sigma and σ_s the width of the two-dimensional Gaussian that describes the spatial emission profile. As this width is also not known, several values are tested for each of the two possibly extended targets: 0 degrees (this is equal to the point-source case), 0.5 degrees, 1 degree, 2 degrees and 5 degrees. In each case, the unbinned maximum likelihood method is used to determine what the best fit for the number of signal events in the data would be.

5.5. Confidence interval and sensitivity calculation

The maximum likelihood method returns the best fit for the number of signal events in the data sample, which will be called $\hat{\mu}$. The next step is to determine the confidence interval for the result. For the analysis described in this thesis, a confidence level of 90% is used, in line with relevant results from the literature.

The confidence interval is determined using the method proposed by Feldman and Cousins [217]. As a start, for each μ_i in a range of possible values for μ (the spacing of these depends on how many significant digits are wanted for the confidence interval), a number of pseudo experiments are performed. For each pseudo experiment, a pseudo data sample is created which consists of all data events, each with a random azimuth value which is different for each pseudo experiment. To this sample, a number of signal events is added. This number is taken randomly from a Poisson distribution with expectation value μ_i, and can be different for each pseudo experiment. The signal events to be added are taken randomly from the signal sample that is being considered (the confidence interval is calculated separately for each annihilation channel, WIMP mass and target).

For each pseudo experiment, maximizing the likelihood results in a value for $\hat{\mu}$, which is the best fit value for μ for this specific pseudo experiment. Then, the rank $R(\mu_i)$ for this pseudo experiment is calculated, which is defined as

$$R(\mu_i) = \frac{\mathcal{L}(\mu_i)}{\mathcal{L}(\hat{\mu})} \tag{5.17}$$

where $\mathcal{L}(\mu_i)$ is the likelihood (which was defined in equation 5.3 for the analysis described in this thesis). By definition, $R(\mu_i)$ is always smaller than or equal to 1, and the larger $R(\mu_i)$ is, the closer μ_i is to being a good fit for the number of signal events in the event sample (since a high $R(\mu_i)$ means that the likelihood for μ is close to that for the best fit $\hat{\mu}$).

Doing this for all pseudo experiments with the same μ_i results in a distribution of $R(\mu_i)$ values, and from this the critical value $R_{crit}(\mu_i)$ is calculated. This is defined as the value that is lower than 90% of the $R(\mu_i)$ distribution. Doing this for each μ_i results in a set of $R_{crit}(\mu_i)$ values (one for each μ_i). Then, one takes the real data sample and calculates $R_{data}(\mu_i)$ for each μ_i. Each μ_i for which $R_{data}(\mu_i) > R_{crit}(\mu_i)$ is then included in the confidence interval.

If the lower bound of the confidence interval is at 0, there is no evidence for a signal, and only an upper limit is given.

The sensitivity is defined as the expected median upper limit in the case that there is no signal. It is determined by doing a set of pseudo experiments without injecting signal events, and for each calculating the upper limit that would result if this pseudo experiment were the real data set. The median of this set of upper limits is the sensitivity, and the one and two sigma bands (the statistical spread on the sensitivity) can be calculated assuming that the upper limits follow a normal distribution.

The sensitivity and limit are calculated in terms of the number of signal events in the total sample. To be able to compare them to other analyses and theories, they need to be calculated in terms of the velocity-averaged dark matter annihilation cross section. This is done using equation 2.3 and taking into account the total livetime of the sample.

5.6. Sensitivities

The final cut on the BDTs described in section 4.4.6 is optimized for the best sensitivity. It is possible to do this separately for each combination of a target, WIMP mass and annihilation channel, but this would be 760 separate optimizations. Instead, the cuts are optimized for each target using a representative sample: 10 TeV WIMPs annihilating to muons for the hard IceCube-dominated BDT, 100 GeV WIMPs annihilating to muons for the soft IceCube-dominated BDT and 100 GeV WIMPs annihilating to b quarks for the DeepCore-dominated BDT. The resulting BDT cuts can be found in table 5.1 for each of the targets described in section 2.5.

The decision whether to use the final-level IceCube-dominated sample resulting from a cut on the hard BDT score or on the soft BDT score is based on which of these gives the best sensitivity. The WIMP masses where the shift from the soft BDT score (for lower WIMP masses) to the hard BDT score (for higher WIMP masses) is made are listed in table 5.2 for all annihilation channels and targets. In figure 5.7, the sensitivities for a range of masses are shown using either a cut on the hard BDT score or a soft BDT score. Here it is clear that indeed the hard BDT performs significantly better for high masses and the soft BDT performs better for lower masses. The relatively small difference between the two sensitivities for the lowest mass (100 GeV) is due to the fact that in this case, the main contribution comes from the DeepCore-dominated sample, which is the same in both cases.

The cuts in table 5.1 were used to calculate the numbers and create the plots shown in section 4.5, for example the final rates in tables 4.6, 4.7 and 4.8. These tables showed that the purity of the DeepCore-dominated final-level sample is higher than that of the IceCube-dominated final-level samples, and the purity of the soft IceCube-dominated samples is higher than that of the hard IceCube-dominated samples. This is because in samples which mainly consist of well-reconstructed events, one can allow more background than in samples which mainly consist of less-well-reconstructed events. The probability that a background event has a space angle (angle between the reconstructed direction and the target) smaller than the estimated error is low, so most background events do not play a significant role in the final likelihood. In that case, it is worthwhile to have looser cuts which keep more signal events.

Target	DC BDT cut	Soft IC BDT cut	Hard IC BDT cut
Combined five dwarfs	0.35	0.275	0.175
Segue 1	0.35	0.275	0.175
Ursa Major II	0.35	0.3	0.225
Willman 1	0.35	0.3	0.175
Coma Berenices	0.35	0.3	0.225
Draco	0.35	0.275	0.225
Virgo cluster	0.35	0.3	0.175
M31 galaxy	0.35	0.275	0.225

Table 5.1.: The BDT cuts which result in the optimal sensitivity for each of the final-level BDTs and each of the targets considered in this analysis. DC refers to the DeepCore-dominated sample and IC to the IceCube-dominated sample.

Target	$\nu\nu$	$\mu\mu$	$\tau\tau$	WW	bb
Combined five dwarfs	400 GeV	600 GeV	600 GeV	400 GeV	1500 GeV
Segue 1	400 GeV	600 GeV	600 GeV	400 GeV	1500 GeV
Ursa Major II	250 GeV	600 GeV	600 GeV	400 GeV	1500 GeV
Willman 1	400 GeV	600 GeV	600 GeV	400 GeV	1500 GeV
Coma Berenices	400 GeV	600 GeV	600 GeV	400 GeV	1500 GeV
Draco	400 GeV	600 GeV	600 GeV	400 GeV	1500 GeV
Virgo cluster	250 GeV	400 GeV	400 GeV	400 GeV	1500 GeV
M31 galaxy	250 GeV	850 GeV	850 GeV	600 GeV	1500 GeV

Table 5.2.: This table shows for each annihilation channel and target at which WIMP mass the switch from the sample selected by cutting on the soft IceCube-dominated BDT score (used for lower WIMP masses) to the sample selected by cutting on the hard IceCube-dominated BDT score (used for higher WIMP masses) is made.

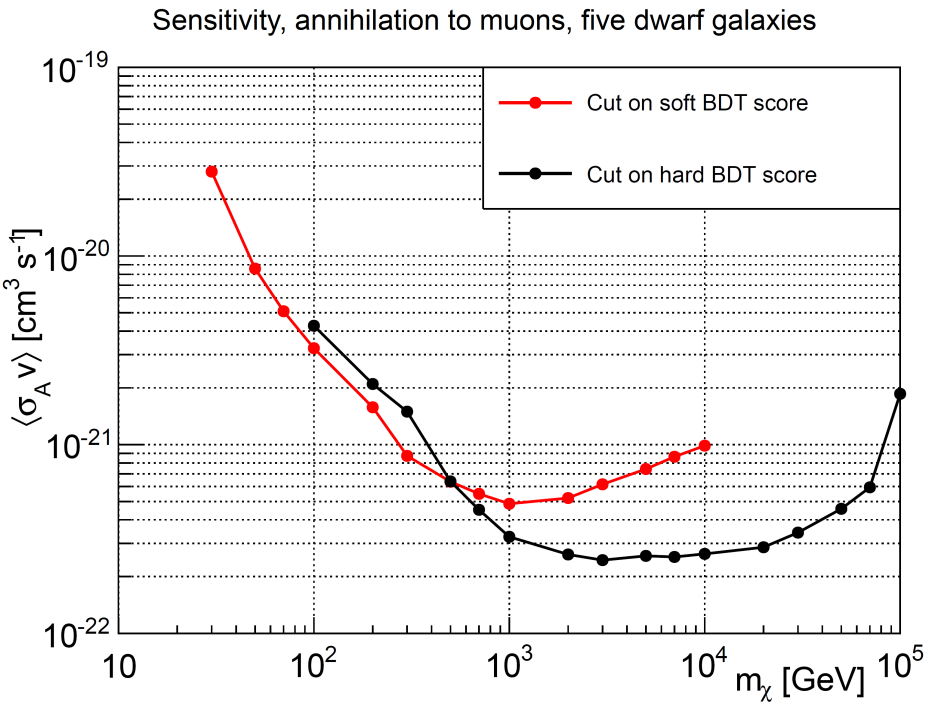


Figure 5.7.: The sensitivity to the velocity-averaged WIMP annihilation cross section for annihilation to muons for the five dwarf galaxies considered in this analysis, using a cut on either the soft BDT score or the hard BDT score to select the final IceCube-dominated sample.

The sensitivities for all targets, WIMP masses and annihilation channels (using the samples after applying the optimal cuts as given in table 5.1) were calculated following the method described in section 5.5. They are shown in the following plots: in figure 5.8 the sensitivities for annihilation to muons are shown for each of the five dwarfs separately and the combined dwarfs. As expected, the sensitivity for the combined dwarfs is significantly better than that for the best single dwarf (Segue 1). The sensitivities do not have the same dependence on WIMP mass because the different dwarfs are located at different zeniths. In figure 5.9, the sensitivities for annihilation to muons for the combined dwarfs, the M31 galaxy and the Virgo cluster are compared. As could already be seen from the J-factors in table 2.1, the sensitivity of the combined dwarfs is significantly better than that of the other two sources. Finally, the sensitivities for different annihilation channels are compared in figures 5.10 (for the combined dwarf galaxies), 5.11 (for the M31 galaxy) and 5.12 (for the Virgo cluster). In general, these follow the behaviour that one would expect from comparing the energy spectra for different channels in figure 2.15: sensitivities are better for channels which produce more (high-energy) neutrinos per WIMP annihilation.

In all cases, the sensitivities are best for WIMP masses between approximately 1 TeV and 10 TeV. At lower masses, the neutrinos resulting from WIMP annihilations have low energies so it is harder to reconstruct their direction and they have a lower probability of passing the event selection. At higher masses, the expected number of WIMPs is lower than at low masses (assuming the same WIMP mass density) so there are less annihilations and thus less signal neutrinos.

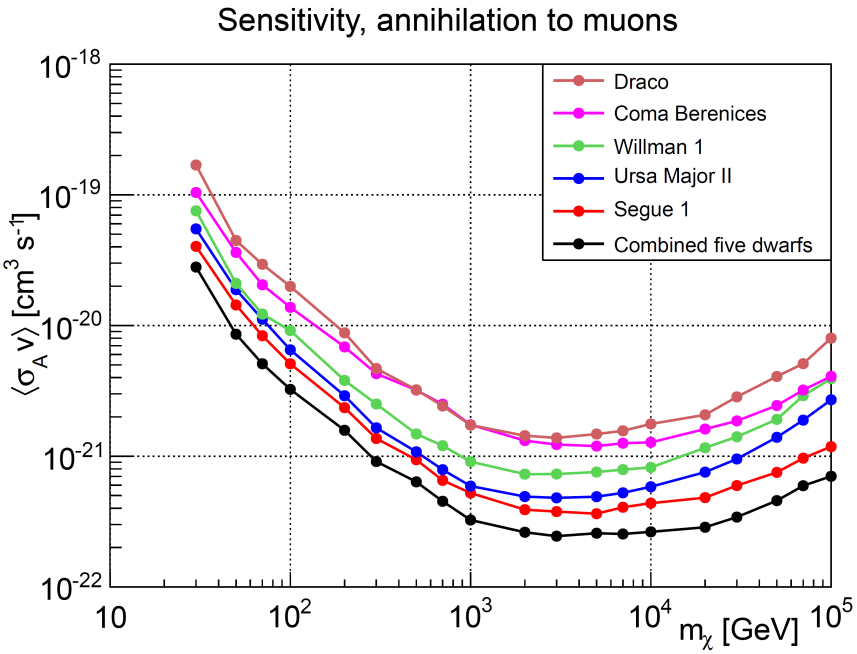


Figure 5.8.: The sensitivity to the velocity-averaged WIMP annihilation cross section for annihilation to muons for the five dwarf galaxies considered in this analysis, both separately and combined.

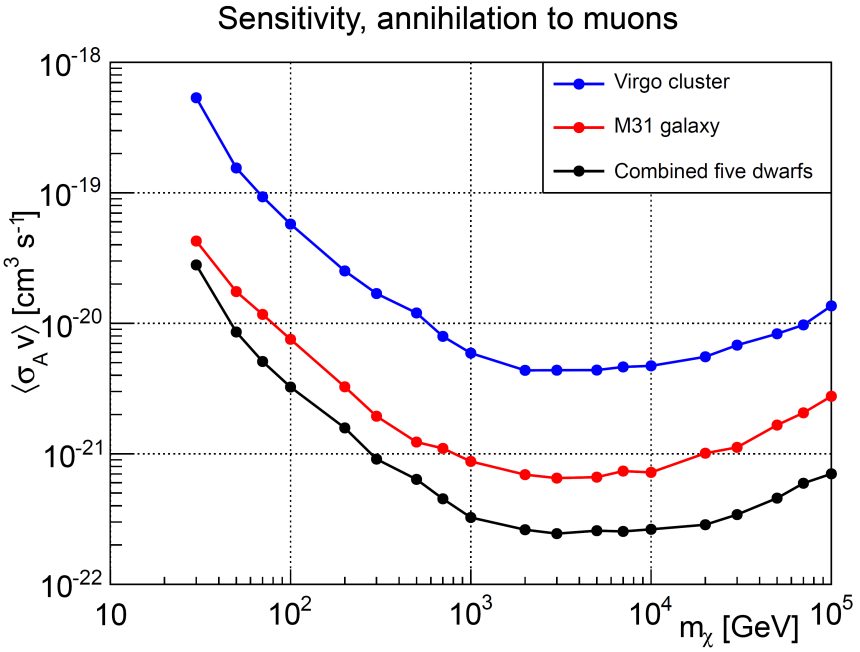


Figure 5.9.: The sensitivity to the velocity-averaged WIMP annihilation cross section for annihilation to muons for the M31 galaxy, the Virgo cluster and the combination of the five dwarf galaxies considered in this analysis.

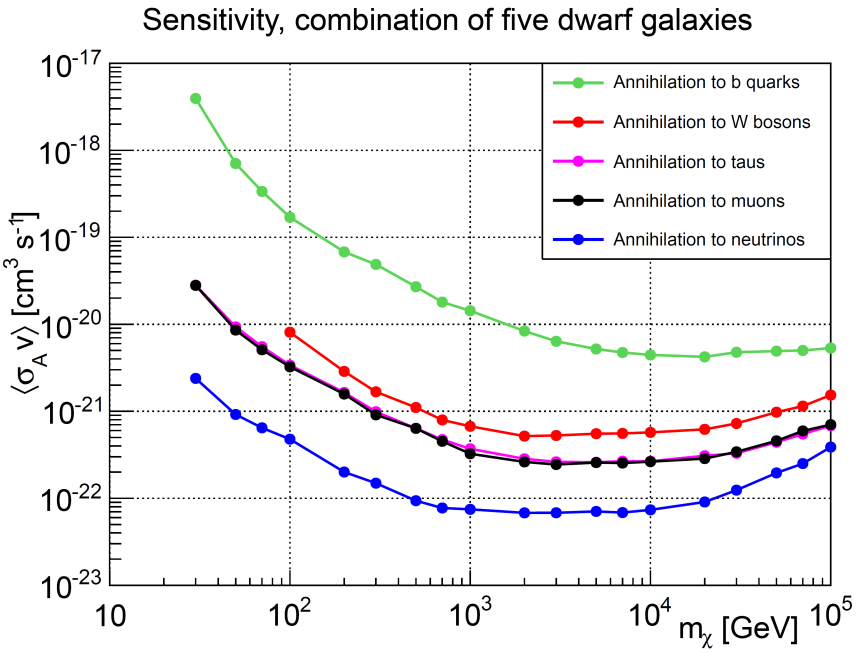


Figure 5.10.: The sensitivity to the velocity-averaged WIMP annihilation cross section for the combination of the five dwarf galaxies considered in this analysis, comparing different annihilation channels.

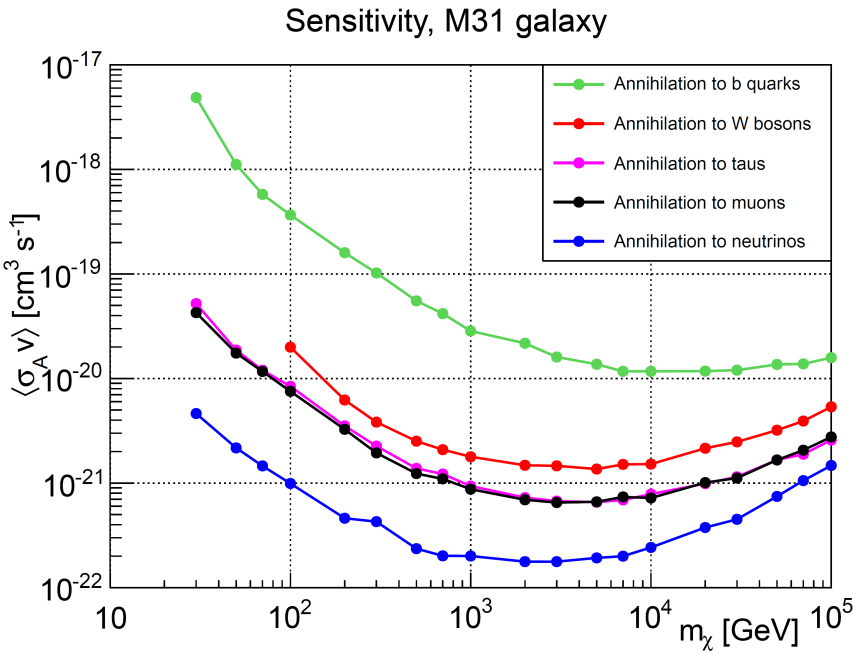


Figure 5.11.: The sensitivity to the velocity-averaged WIMP annihilation cross section for the M31 galaxy, comparing different annihilation channels.

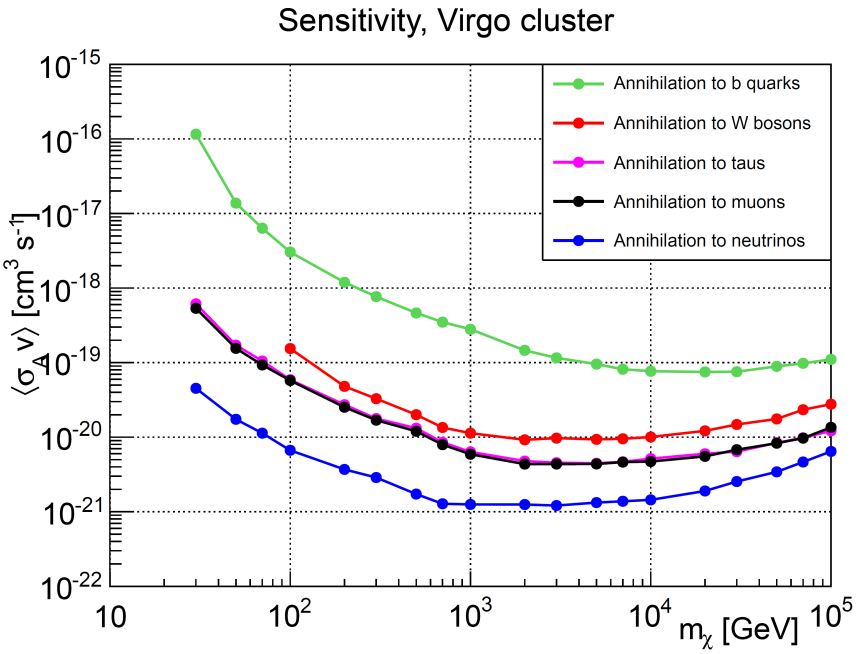


Figure 5.12.: The sensitivity to the velocity-averaged WIMP annihilation cross section for the Virgo cluster, comparing different annihilation channels.

6. Results

The final step of the analysis described in this thesis is to apply the analysis method discussed in chapter 5 to the final level sample resulting from the event selection described in section 4.4. The results of this final step are shown in sections 6.1 and 6.2. In addition to the statistical uncertainties which were discussed in chapter 5, there are also systematic uncertainties which need to be taken into account, these are discussed in section 6.3. Finally, the results from this analysis are compared to those of other analyses in section 6.4.

As is shown in section 6.4, the results from this analysis are up to an order of magnitude more stringent than those from the previous IceCube analysis studying the same targets. This is due to a number of improvements, both general and specifically developed for this analysis, which have been discussed in detail in the previous chapters and are summarized in section 6.4.

6.1. Limits

Five annihilation channels, nineteen WIMP masses and eight targets (the M31 galaxy, the Virgo cluster, and the five dwarf galaxies separately and combined) are studied in this analysis. For each of these, the best-fit number of signal events and the 90% C.L. upper and lower limits on this number were calculated following the methodology described in sections 5.1 and 5.5. From this, the 90% C.L. upper limit on the velocity-averaged WIMP annihilation cross section was calculated using equation 2.3. This upper limit was compared to the sensitivity which was calculated in section 5.6. As an example, the results for annihilation to muons are shown in tables 6.1 (combination of five dwarfs), 6.2 (M31 galaxy) and 6.3 (Virgo cluster). For the five dwarfs separately, these tables can be found in appendix D. For the other annihilation channels, the overall behaviour is similar to the behaviour for annihilation to muons, except that under- or overfluctuations will show up at higher masses for annihilation to b quarks and W bosons and at lower masses for direct annihilation to neutrinos. In all results shown in this section, systematic uncertainties have not yet been taken into account.

In figures 6.1, 6.2 and 6.3 these results are shown visually, as the 90% C.L. limit on the velocity-averaged WIMP annihilation cross section is compared to the sensitivity and the statistical spread on the sensitivity. From the figures and tables one can see that while the limit is close to the sensitivity for the M31 galaxy and there is a small underfluctuation for the low masses for the Virgo cluster, there is also an overfluctuation for the high masses for the combination of the five dwarfs. This points towards either a signal, an error in the method or a non-significant statistical overfluctuation, so it needs to be treated with care.

The first step is to test how likely it is to detect such an overfluctuation in the case that the dataset consists only of background events. For this, one needs to take into account that the probability of detecting an overfluctuation gets larger when more combinations of WIMP masses and annihilation channels are being tested. This is usually referred to as the

m_χ [GeV]	$\hat{\mu}$	$\mu_{l,90\%}$	$\mu_{h,90\%}$	$\langle\sigma_A v\rangle_{sens}$ [10^{-21} cm^3 s^{-1}]	$\langle\sigma_A v\rangle_{lim}$ [10^{-21} cm^3 s^{-1}]
30	0.0	0.0	18.5	28	25
50	0.0	0.0	22.0	8.6	7.5
70	2.7	0.0	31.0	5.1	5.6
100	9.7	0.0	48.0	3.2	4.5
200	37.7	0.0	90.0	1.6	2.6
300	42.2	8.8	90.0	0.91	1.7
500	38.2	9.8	78.0	0.64	1.2
700	10.4	0.0	44.0	0.45	0.55
1000	9.1	0.0	39.0	0.32	0.42
2000	6.7	0.0	34.0	0.26	0.34
3000	5.8	0.0	32.0	0.24	0.34
5000	6.9	0.0	30.0	0.26	0.37
7000	7.1	0.0	28.0	0.25	0.40
10000	9.1	0.1	28.0	0.26	0.48
20000	10.1	2.7	26.0	0.29	0.71
30000	9.8	2.9	23.0	0.34	0.87
50000	8.9	2.8	21.0	0.46	1.3
70000	8.4	2.6	20.0	0.59	1.7
100000	7.7	2.4	18.0	0.70	2.3

Table 6.1.: The best fit ($\hat{\mu}$) and 90% C.L. lower and upper limits ($\mu_{l,90\%}$ and $\mu_{h,90\%}$) for the number of signal events μ in the total dataset, and the sensitivity to and 90% C.L. upper limit on the velocity-averaged annihilation cross section $\langle\sigma_A v\rangle$, for different WIMP masses m_χ. The results are shown for annihilation to muons and the combination of the five dwarf galaxies (Segue 1, Ursa Major II, Willman 1, Coma Berenices and Draco). Systematic uncertainties have not yet been taken into account for these limits.

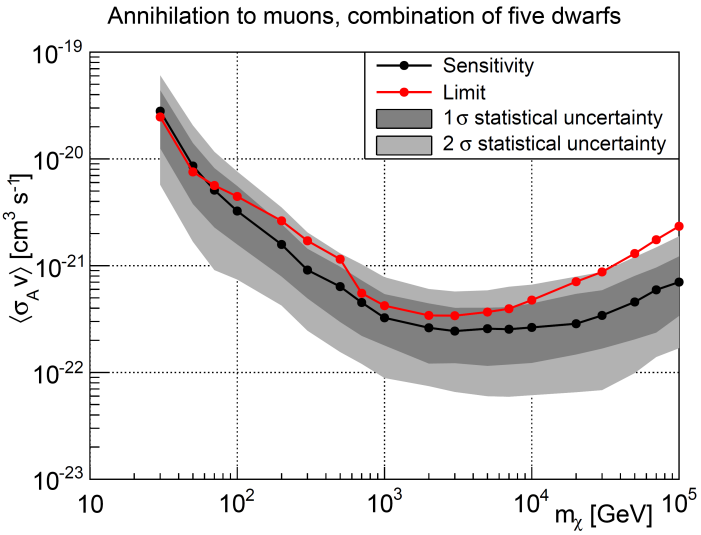


Figure 6.1.: Comparing the sensitivity (black) and 90% C.L. limit (red) on the velocity-averaged WIMP annihilation cross section for annihilation to muons for the combination of five dwarf galaxies (Segue 1, Ursa Major II, Willman 1, Coma Berenices and Draco). The statistical spread on the sensitivity is also shown (grey one- and two-sigma bands). The markers denote which WIMP masses have been tested, the lines have been added to guide the eye. Systematic uncertainties have not yet been taken into account for these limits.

m_χ [GeV]	$\hat{\mu}$	$\mu_{l,90\%}$	$\mu_{h,90\%}$	$\langle\sigma_A v\rangle_{lim}$ [10^{-21} cm^3 s^{-1}]	$\langle\sigma_A v\rangle_{sens}$ [10^{-21} cm^3 s^{-1}]
30	2.9	0.0	17.5	43	68
50	3.3	0.0	22.0	17	24
70	0.0	0.0	21.0	12	12
100	0.0	0.0	25.0	7.5	7.5
200	0.0	0.0	30.0	3.3	3.3
300	0.0	0.0	24.0	1.9	1.7
500	0.0	0.0	19.0	1.2	1.1
700	0.66	0.0	17.5	1.1	0.96
1000	0.0	0.0	11.5	0.87	0.77
2000	0.14	0.0	10.5	0.69	0.63
3000	0.14	0.0	10.5	0.65	0.65
5000	0.0	0.0	8.8	0.66	0.62
7000	0.0	0.0	8.2	0.74	0.66
10000	0.0	0.0	7.4	0.72	0.70
20000	0.0	0.0	6.2	1.0	0.92
30000	0.0	0.0	5.4	1.1	1.1
50000	0.0	0.0	5.0	1.7	1.7
70000	0.0	0.0	4.8	2.1	2.3
100000	0.0	0.0	4.4	2.8	3.1

Table 6.2.: The best fit and 90% C.L. lower and upper limits for the number of signal events μ in the total dataset, and the sensitivity to and 90% C.L. upper limit on the velocity-averaged annihilation cross section $\langle\sigma_A v\rangle$, for different WIMP masses m_χ. The results are shown for annihilation to muons and the M31 galaxy. Systematic uncertainties have not yet been taken into account for these limits.

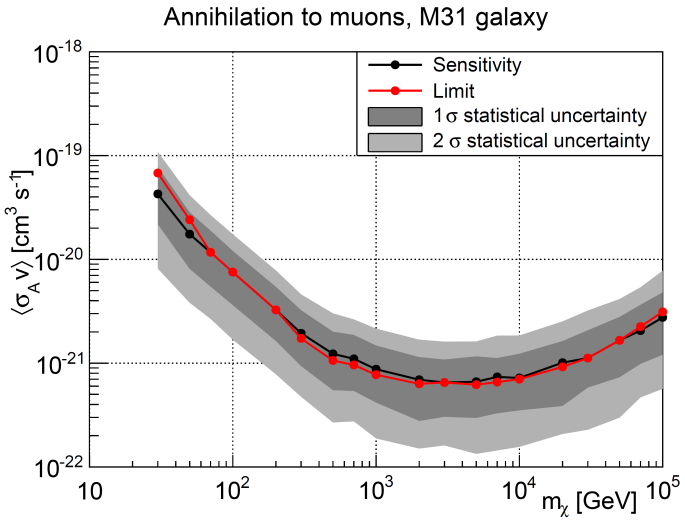


Figure 6.2.: Comparing the sensitivity (black) and 90% C.L. limit (red) on the velocity-averaged WIMP annihilation cross section for annihilation to muons for the M31 galaxy. The statistical spread on the sensitivity is also shown (grey one- and two-sigma bands). The markers denote which WIMP masses have been tested, the lines have been added to guide the eye. Systematic uncertainties have not yet been taken into account for these limits.

m_χ [GeV]	$\hat{\mu}$	$\mu_{l,90\%}$	$\mu_{h,90\%}$	$\langle\sigma_A v\rangle_{lim}$ [10^{-21} cm^3 s^{-1}]	$\langle\sigma_A v\rangle_{sens}$ [10^{-21} cm^3 s^{-1}]
30	0.0	0.0	8.2	$5.3 \cdot 10^2$	$3.2 \cdot 10^2$
50	0.0	0.0	6.2	$1.6 \cdot 10^2$	54
70	0.0	0.0	8.4	93	39
100	0.0	0.0	12.5	58	34
200	0.0	0.0	15.0	25	18
300	0.0	0.0	11.0	17	10
500	0.0	0.0	9.0	12	6.4
700	0.0	0.0	10.5	7.9	5.4
1000	0.0	0.0	10.5	5.9	4.3
2000	0.0	0.0	12.0	4.4	4.0
3000	0.0	0.0	14.0	4.4	4.7
5000	0.0	0.0	14.5	4.4	5.3
7000	0.02	0.0	13.5	4.6	5.4
10000	0.0	0.0	12.0	4.7	5.7
20000	0.0	0.0	9.0	5.5	6.4
30000	0.0	0.0	8.4	6.8	7.9
50000	0.0	0.0	7.0	8.3	10
70000	0.0	0.0	6.4	9.7	12
100000	0.0	0.0	5.4	14	15

Table 6.3.: The best fit and 90% C.L. lower and upper limits for the number of signal events μ in the total dataset, and the sensitivity to and 90% C.L. upper limit on the velocity-averaged annihilation cross section $\langle\sigma_A v\rangle$, for different WIMP masses m_χ. The results are shown for annihilation to muons and the Virgo cluster. Systematic uncertainties have not yet been taken into account for these limits.

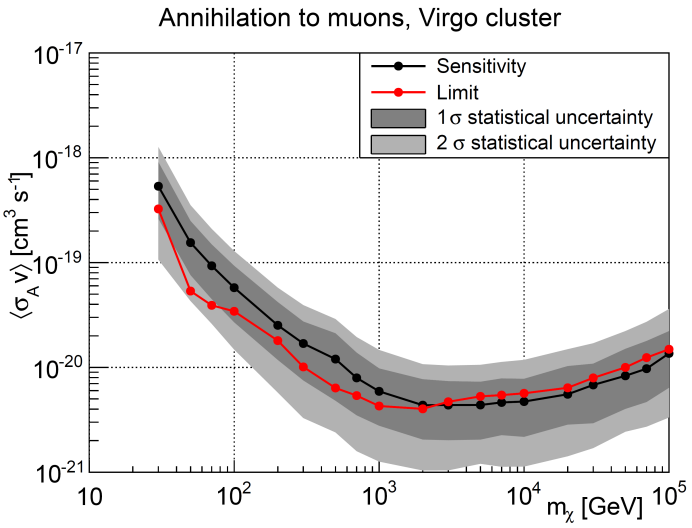


Figure 6.3.: Comparing the sensitivity (black) and 90% C.L. limit (red) on the velocity-averaged WIMP annihilation cross section for annihilation to muons for the Virgo cluster. The statistical spread on the sensitivity is also shown (grey one- and two-sigma bands). The markers denote which WIMP masses have been tested, the lines have been added to guide the eye. Systematic uncertainties have not yet been taken into account for these limits.

trial factor. If the experiments were all independent, it would be a matter of multiplying the probability of detecting an overfluctuation in one 'experiment' (one WIMP mass and annihilation channel) with the number of experiments. However, in this case the experiments are not independent.

To accurately determine the probability of detecting an overfluctuation of some size, assuming that the data consists only of background events, one first needs a measure for the size of an overfluctuation, which can be used to compare different experiments. For this, the test statistic TS is used, which is defined as

$$TS = -2 \cdot \log\left(\frac{\mathcal{L}(0)}{\mathcal{L}(\hat{\mu})}\right) \tag{6.1}$$

with $\mathcal{L}(\hat{\mu})$ the likelihood for the best-fit μ for this particular sample (defined in equation 5.3) and $\mathcal{L}(0)$ the likelihood for the background-only assumption. Since $\mathcal{L}(\hat{\mu}) \geq \mathcal{L}(0)$, $TS \geq 0$, and the more TS deviates from 0, the larger the overfluctuation (since $\hat{\mu}$ is constrained to be non-negative, underfluctuations will result in a TS of 0).

The method applied here is to first calculate the TS for each WIMP mass and annihilation channel, and then determine which of these has the highest TS. For the combination of the five dwarf galaxies, the highest TS is 6.71, for 30 TeV WIMPs annihilating directly to neutrinos. Then, a large number of background-only pseudo data samples is created, and each of these are treated like they are data: all WIMP masses and channels are tested, and the largest TS out of all of these is determined. This results in a distribution of highest TS values for background-only pseudo experiments, which is shown in figure 6.4. The probability of detecting an overfluctuation as large as the one that is detected in data or larger is then the fraction of the distribution of highest TS values that is equal to or larger than 6.71, which is 4.9%. This means that while there is a clear overfluctuation, it is not significant enough to claim detection of a signal, so it will be assumed that this is a statistical overfluctuation and only upper limits on the velocity-averaged WIMP annihilation cross section will be quoted.

When considering the results for the single dwarfs as shown in appendix D, one can see that the overfluctuation for the combination of the five dwarfs is driven by overfluctuations at high WIMP masses for Ursa Major II, Willman 1 and Coma Berenices. Such an overfluctuation does not appear for Segue 1 and Draco. Since Draco has a significantly lower J-factor than the other dwarfs, if the fitted number of signal events for the combination of five dwarfs (which is approximately 10) is true, the expected number of signal events from Draco would be less than 1. For Segue this is not the case, but if the overfluctuation were due to signal, it could be that the J-factor estimate for Segue is not accurate or that there is an underfluctuation for Segue.

For this analysis three years of data are combined, so it is also interesting to consider whether one of these years contributes significantly more to the overfluctuation than the other years, as this might point towards issues with the data. In figure 6.5, the sensitivities and limits for the combination of the five dwarf galaxies are compared for the three years separately, for WIMP masses above 10 TeV. There is an overfluctuation for all three years, which is largest for 2011 but also at the one-sigma level for the other two years. This means that there is no evidence for issues with the data of any particular year.

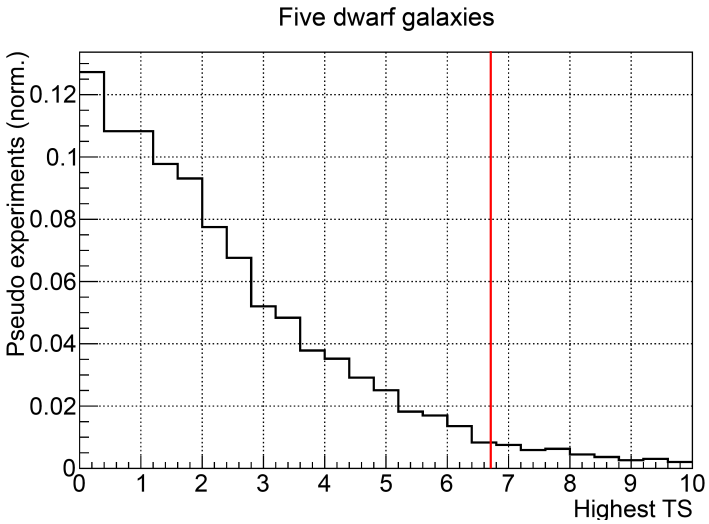
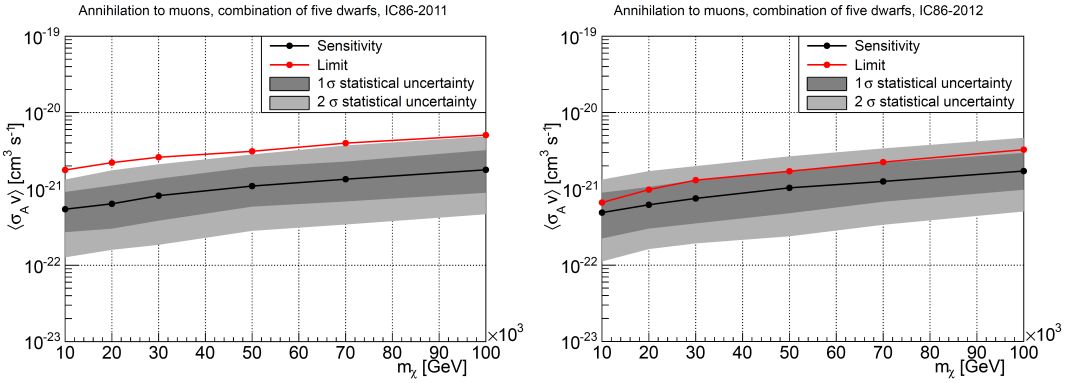
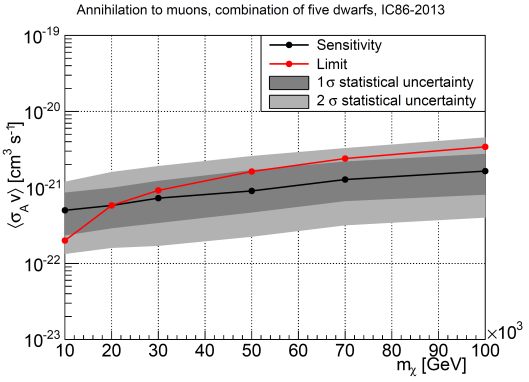


Figure 6.4.: The distribution of highest test statistic values (out of all tested WIMP masses and annihilation channels) for background-only pseudo experiments for the combination of five dwarfs. The red line denotes the highest TS value for the experimental dataset, which is 6.71. 4.9% of the distribution is above this value.



(a) IC86-2011

(b) IC86-2012



(c) IC86-2013

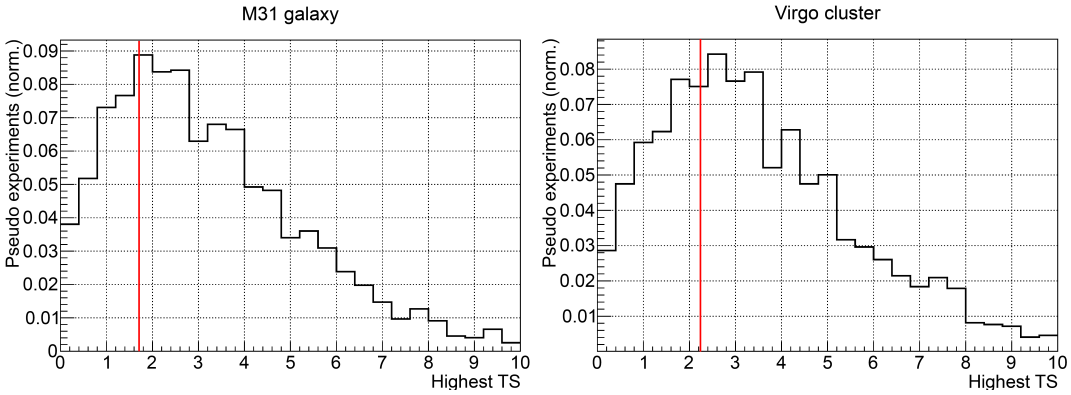
Figure 6.5.: Comparing the sensitivity (black) and 90% C.L. limit (red) on the velocity-averaged WIMP annihilation cross section for the combination of the five dwarf galaxies, for each of the three years of data separately. The statistical spread on the sensitivity is also shown (grey one- and two-sigma bands). The markers denote which WIMP masses have been tested, the lines have been added to guide the eye. Systematic uncertainties have not yet been taken into account for these limits.

In conclusion, there is an overfluctuation at high WIMP masses for the combination of the five dwarf galaxies considered in this analysis. There are no signs that this is due to an error in the method. Since the probability of getting an overfluctuation of this size or larger is 4.9%, it will be treated as a statistical overfluctuation. For the M31 galaxy and the Virgo cluster, there is no sign of an overfluctuation if they are treated as point sources. The next section will discuss the search for an extended signal from these two targets.

6.2. Extended targets

As described in section 5.4, it is possible to use the unbinned maximum likelihood method to search for an extended signal by adapting the spatial signal PDF. For the analysis described in this thesis, the extended signal is assumed to follow a two-dimensional Gaussian distribution, and widths of 0, 0.5, 1, 2 and 5 degrees are tested.

To determine whether a significant signal has been detected, a method is applied that is similar to what was used for the combined dwarf galaxies. For each of the two targets (the M31 galaxy and the Virgo cluster), the test statistic TS (defined in equation 6.1) is calculated for each WIMP mass, annihilation channel and distribution width. For the M31 galaxy, the highest TS is 1.71, for 30 TeV WIMPs annihilating to neutrinos and a width of 5 degrees. For the Virgo cluster, the highest TS is 2.24, for 7 TeV WIMPs annihilating to neutrinos and a width of 2 degrees.



(a) M31 galaxy **(b)** Virgo cluster

Figure 6.6.: The distribution of highest test statistic values (out of all tested WIMP masses and annihilation channels) for background-only pseudo experiments for the M31 galaxy and the Virgo cluster. The red lines denote the highest TS value for the experimental dataset, which is at 1.71 for the M31 galaxy and at 2.24 for the Virgo cluster. In both cases, more than 50% of the distribution is above this value.

Like for the combined dwarf galaxies, for these targets a number of background-only pseudo datasets are created and for each of these, the TS values are calculated for each combination of a WIMP mass, an annihilation channel and a distribution width. The resulting distributions of highest TS values are shown in figure 6.6, and the probability of getting the same or a larger overfluctuation is over 50% for both the M31 galaxy and the Virgo cluster. This means that there is no sign of an extended emission from the M31 galaxy or the Virgo cluster.

6.3. Systematic uncertainties

In the limits shown in section 6.1, only statistical uncertainties have been taken into account. In this section, systematic uncertainties will be included. Since the background for this analysis is taken from data, the only systematic uncertainty related to background is that due to preferred directions in the detector. As the detector is not symmetrical, these do exist, but due to the rotation of the Earth their eventual effect is negligible. This means that only systematic uncertainties related to the signal simulation need to be taken into account. The main ones are those related to the detector simulation: the uncertainties on the ice model (both for the absorption and scattering coefficients) and on the optical efficiency of the DOMs.

In this analysis, systematic uncertainties are taken into account in the following way: signal simulation datasets have been produced where the absorption length, the scattering length or the opticial efficiency have been decreased by 10%. These datasets have been used to calculate the limit on the velocity-averaged WIMP annihilation cross section assuming that this is the true signal simulation. This means that signal injection is done using these so-called 'systematic datasets', but for the likelihood calculation the PDFs based on the baseline simulation are used. The resulting percentual systematic uncertainties on the limits are added in quadrature and the limits are worsened by this percentage, which is different for each target, WIMP mass and annihilation channel.

In the final stages of the analysis, the high-energy (NuGen) ice model systematic datasets were discovered to have been simulated incorrectly. From earlier studies it was clear that the effect of changes in the ice model is larger for lower WIMP masses, so it was decided to use the (percentual) systematic effect on the 100 GeV WIMPs (the highest mass where the NuGen dataset is not used) for higher-mass WIMPs as well. The resulting systematic effects on the limits for WIMPs annihilating to muons for the combined five dwarf galaxies are shown in table 6.4. Also, the limits including systematic effects are compared to the limits without systematic effects in figure 6.7 for the combined five dwarf galaxies, the M31 galaxy and the Virgo cluster.

To complete the set of limits with systematic uncertainties included, figures 6.8 and 6.9 show the limits including systematic uncertainties for each of the dwarf galaxies separately (compared to the five combined dwarf galaxies) for annihilation to muons, and for each channel that is considered for the combination of the five dwarf galaxies. These figures can be compared to figures 5.8 and 5.10, which show the sensitivities for these two cases. From this comparison one can see that while the sensitivities scale inversely with the J-factors of the dwarf galaxies and the sensitivity for the combination of the five dwarf galaxies is better than that for each of the dwarf galaxies separately, this is not always the case for the limits. This is caused by the statistical under- and overfluctuations, for example three of the dwarf galaxies had clear overfluctuations at high WIMP masses (as discussed above) and the Willman 1 dwarf galaxy had a large underfluctuation for relatively low masses. These under- and overfluctuations are also visible in the plots in appendix D.

6.4. Comparison to other limits

As has been discussed in section 2.3.3, many other indirect searches for annihilating dark matter have been performed, both within the IceCube collaboration and in other experi-

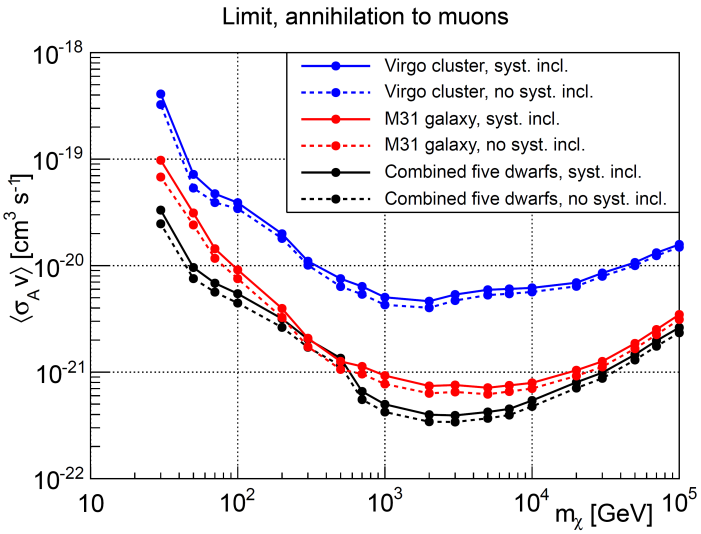


Figure 6.7.: Comparing the limits on the velocity-averaged WIMP annihilation cross section before and after including systematic uncertainties, for the combination of five dwarf galaxies, the M31 galaxy and the Virgo cluster. The markers denote which WIMP masses have been tested, the lines have been added to guide the eye.

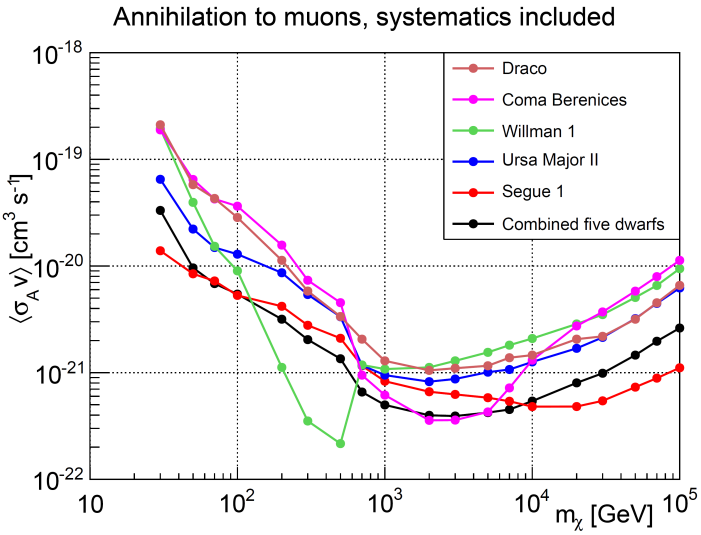


Figure 6.8.: Comparing the limits (including systematic uncertainties) on the velocity-averaged WIMP annihilation cross section from the combination of five dwarf galaxies and each dwarf galaxy separately, for annihilation to muons. The markers denote which WIMP masses have been tested, the lines have been added to guide the eye.

m_χ [GeV]	$\langle\sigma_A v\rangle_{no}$ [cm^3 s^{-1}]	DOM eff.	Abs.	Scat.	Total	$\langle\sigma_A v\rangle_{sys}$ [cm^3 s^{-1}]
30	$25 \cdot 10^{-21}$	33.4%	10.1%	15.6%	38.2%	$34 \cdot 10^{-21}$
50	$7.5 \cdot 10^{-21}$	21.3%	9.9%	14.3%	27.5%	$9.6 \cdot 10^{-21}$
70	$5.6 \cdot 10^{-21}$	17.1%	9.8%	11.9%	23.0%	$6.9 \cdot 10^{-21}$
100	$4.5 \cdot 10^{-21}$	16.9%	9.5%	11.1%	22.4%	$5.5 \cdot 10^{-21}$
200	$2.6 \cdot 10^{-21}$	14.2%	9.5%	11.1%	20.4%	$3.2 \cdot 10^{-21}$
300	$1.7 \cdot 10^{-21}$	12.6%	9.5%	11.1%	19.4%	$2.0 \cdot 10^{-21}$
500	$1.2 \cdot 10^{-21}$	9.3%	9.5%	11.1%	17.3%	$1.4 \cdot 10^{-21}$
700	$0.55 \cdot 10^{-21}$	14.8%	7.3%	9.7%	19.1%	$0.66 \cdot 10^{-21}$
1000	$0.42 \cdot 10^{-21}$	13.4%	7.3%	9.7%	18.1%	$0.50 \cdot 10^{-21}$
2000	$0.34 \cdot 10^{-21}$	10.8%	7.3%	9.7%	16.2%	$0.40 \cdot 10^{-21}$
3000	$0.34 \cdot 10^{-21}$	9.5%	7.3%	9.7%	15.4%	$0.39 \cdot 10^{-21}$
5000	$0.37 \cdot 10^{-21}$	8.3%	7.3%	9.7%	14.7%	$0.42 \cdot 10^{-21}$
7000	$0.40 \cdot 10^{-21}$	7.6%	7.3%	9.7%	14.3%	$0.45 \cdot 10^{-21}$
10000	$0.48 \cdot 10^{-21}$	6.2%	7.3%	9.7%	13.6%	$0.54 \cdot 10^{-21}$
20000	$0.71 \cdot 10^{-21}$	5.3%	7.3%	9.7%	13.2%	$0.80 \cdot 10^{-21}$
30000	$0.87 \cdot 10^{-21}$	4.6%	7.3%	9.7%	13.0%	$0.99 \cdot 10^{-21}$
50000	$1.3 \cdot 10^{-21}$	3.6%	7.3%	9.7%	12.6%	$1.5 \cdot 10^{-21}$
70000	$1.7 \cdot 10^{-21}$	3.3%	7.3%	9.7%	12.6%	$2.0 \cdot 10^{-21}$
100000	$2.3 \cdot 10^{-21}$	1.0%	7.3%	9.7%	12.2%	$2.6 \cdot 10^{-21}$

Table 6.4.: The 90% C.L. upper limit on the velocity-averaged annihilation cross section $\langle\sigma_A v\rangle$ without including systematic uncertainties, the percentual systematic uncertainties due to uncertainties on absorption in the ice, scattering in the ice and the DOM optical efficiency and the squared sum of these, and the 90% C.L. upper limit including systematic uncertainties. The uncertainty due to absorption in the ice and scattering in the ice for WIMP masses above 100 GeV is calculated using the sample for 100 GeV. The change between 500 and 700 GeV is due to the switch from using the soft IceCube-dominated BDT to using the hard IceCube-dominated BDT. This table is for annihilation to muons and using the combination of the five dwarf galaxies (Segue 1, Ursa Major II, Willman 1, Coma Berenices and Draco) as a target.

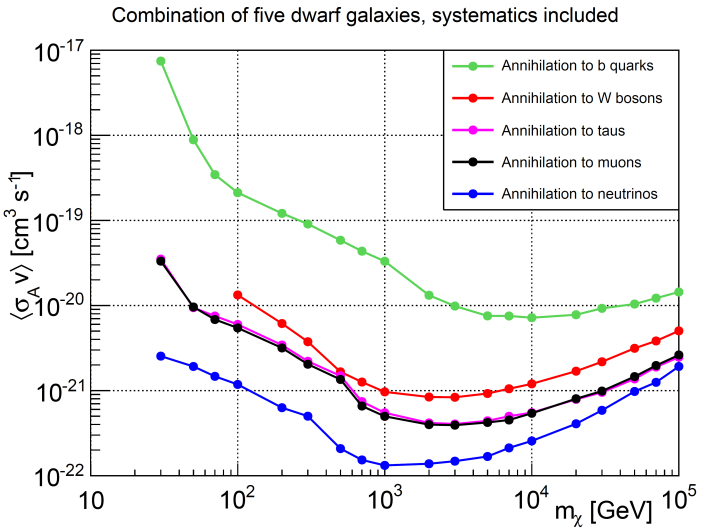


Figure 6.9.: Comparing the limits (including systematic uncertainties) on the velocity-averaged WIMP annihilation cross section from the combination of five dwarf galaxies for different annihilation channels. The markers denote which WIMP masses have been tested, the lines have been added to guide the eye.

ments. In this section, the results of the analysis described in this thesis are compared to other limits on the velocity-averaged WIMP annihilation cross section.

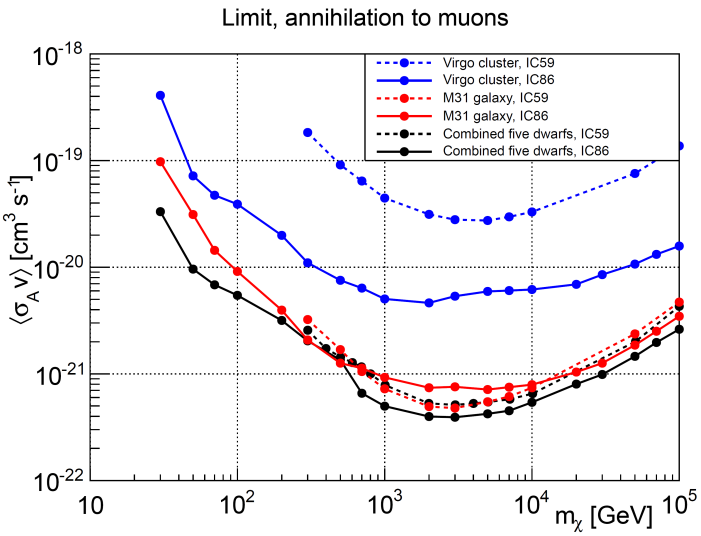


Figure 6.10.: Comparing the limits (including systematic uncertainties) on the velocity-averaged WIMP annihilation cross section for the IceCube-59 analysis and the IceCube-86 analysis, for the combination of dwarf galaxies, the M31 galaxy and the Virgo cluster.

First, the results of the analysis described in this thesis are compared to the previous IceCube analysis that considered (almost) the same targets. This analysis used data from an incomplete version of the IceCube detector (IC59, containing 59 strings instead of 86). In figure 6.10, the results from these two analyses are compared for the combination of dwarf galaxies, the M31 galaxy and the Virgo cluster. One difference between the two analyses

is that while the IC59 analysis calculated limits for WIMP masses down to 300 GeV, the analysis described in this thesis goes down another order of magnitude in mass, to 30 GeV. This is possible due to the existence of the DeepCore subdetector (see section 3.3.1), which did not yet exist in IC59 and which is capable of detecting neutrinos down to energies of the order of 10 GeV.

One can also see that the improvement in the limit is not a constant factor, but depends on the target and WIMP mass. This is mainly due to statistical under- and overfluctuations. For example, in the case of the M31 galaxy, there was a large underfluctuation in the IC59 analysis, so for some masses the IC59 limit is better than the IC86 limit. In general, the limits on the velocity-averaged annihilation cross section are significantly better for the analysis described here than for the IC59 analysis, which is due to the following improvements:

- The IC59 analysis used a binned analysis method, where a circular region around the target is defined as the signal region and all events inside this region are counted. The IC86 analysis used an improved analysis technique, described in section 5.1, which takes into account how large the angle between the reconstructed direction and the target direction is, what the expected error of the reconstruction is, and what the reconstructed energy proxy of the event is.

- The IC86 analysis used a different event selection, which was optimized specifically for neutrinos from WIMP annihilations and a large range of WIMP masses.

- The IC86 analysis used an improved direction reconstruction. The SplineMPE direction reconstruction was used, while the IC59 analysis used the PandelMPE direction reconstruction. As can be seen from figure 4.3, SplineMPE performs better than PandelMPE for all neutrino energies.

- The IC86 analysis used three years of data, compared to one year of data for the IC59 analysis.

- The IC86 analysis used a larger detector, with 86 active strings instead of 59. The IC86 detector also contained the DeepCore subdetector, which especially improved the limits for lower WIMP masses.

- In the case of the combined dwarf galaxies, the IC86 analysis considered five dwarf galaxies while the IC59 analysis considered two dwarf galaxies.

The comparison between the IC59 and IC86 limits is also influenced by the fact that both the J-factors and the Monte Carlo simulation used for the analyses were different. Based on current knowledge, the estimated J-factors for the Segue 1 and Ursa Major II dwarf galaxies are lower than the values that were used for the IC59 analysis, while the estimated J-factor for the Virgo cluster is higher. This is shown in table 6.5. As the limit on the velocity-averaged annihilation cross section is inversely proportional to the assumed J-factor, this has a significant effect. Also, the IC86 analysis used an improved version of the Monte Carlo simulation software compared to the IC59 analysis, which resulted in less simulated neutrinos being detected, which leads to a worse limit. In this aspect, the limits resulting from the IC59 analysis were likely too optimistic.

Target	$\log_{10}(J$ / GeV^2 cm$^{-5})_{IC59}$	$\log_{10}(J$ / GeV^2 cm$^{-5})_{IC86}$
Segue 1	19.6	19.5
Ursa Major II	19.6	19.3
Willman 1	-	19.1
Coma Berenices	19.0	19.0
Draco	18.8	18.8
M31	19.2	19.2
Virgo	18.2	18.5

Table 6.5.: Comparing the logarithm of the J-factors that were used to calculate the limits for the IC59 and IC86 analyses. The Willman 1 dwarf galaxy was not used in the IC59 analysis, so there was no J-factor assumed. For the IC59 analysis, the J-factors were taken from [218] (M31 galaxy), [219] (dwarf galaxies) and [220] (Virgo cluster). For the IC86 analysis, they were taken from [127] (dwarf galaxies), [151] (M31 galaxy) and [152] (Virgo cluster, this is the published version of [220]).

In figure 6.11, the limits from the analysis described in this thesis are compared to limits from two gamma-ray experiments, Fermi and MAGIC, which also considered WIMP annihilation in dwarf galaxies. A similar figure was also shown in chapter 2 (figure 2.10), but the result from the IC59 analysis has now been replaced by the result from the analysis described in this thesis. Since Fermi and MAGIC showed limits up to WIMP masses of 10 TeV it is not possible to compare the limits for higher masses, but for lower masses the limits from the gamma-ray experiments are an order of magnitude or more better than the limits from this analysis. This is due to the low interaction cross section for neutrinos compared to gamma rays (which at least for relatively low-energy neutrinos is not cancelled out by the large size of the IceCube detector compared to for example the Fermi detector) and the large median angular error compared to the typical angular error for gamma-ray experiments. However, gamma-ray experiments so far have not published limits on direct annihilation to neutrinos, which is the best channel for detection with neutrino experiments and one of the worst for detection with gamma-ray experiments.

In figure 6.12, the limits from this analysis are compared to limits from other neutrino searches using different targets: the Galactic Centre and the Galactic Halo. Also in this case, one can conclude that the limits from other searches are significantly better than those from the analysis described in this thesis for all WIMP masses. This is mainly due to the much larger J-factor of the Galactic Centre and Halo compared to the dwarf galaxies, the M31 galaxy and the Virgo cluster.

However, the different analyses shown in this section all have different uncertainties. For example, the Galactic Centre analysis is very dependent on the dark matter halo profile, which is not the case for the dwarf galaxy analysis. This means that it is still worthwhile and interesting to consider different targets, as signals which show up in one analysis but not in the other could signify that certain assumptions are incorrect. As there is still so little known about dark matter, one cannot know for certain where it will show up for the first time, but dwarf galaxies remain a good candidate.

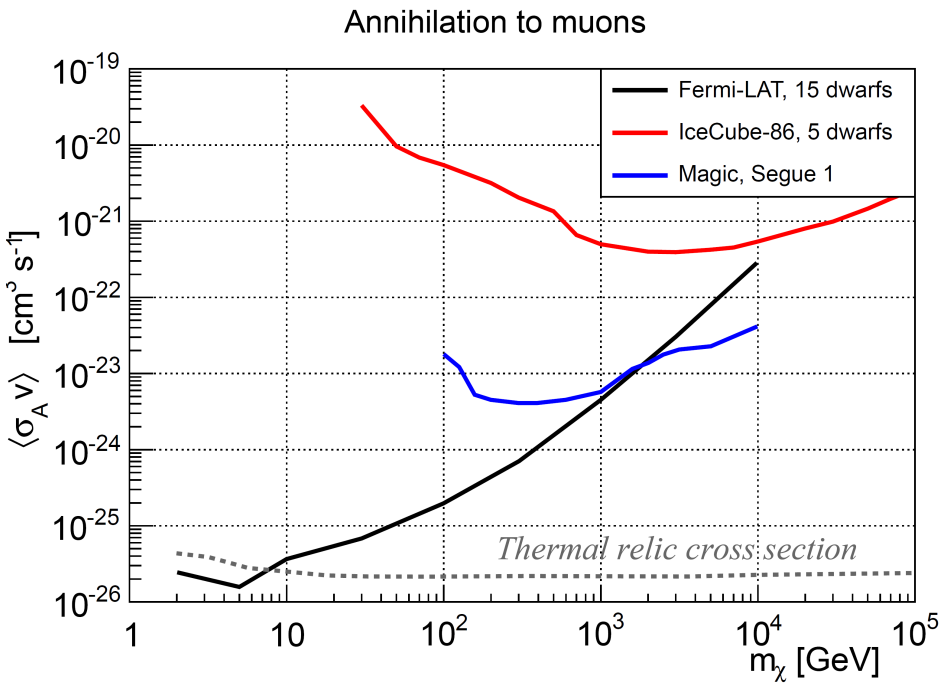


Figure 6.11.: Comparison of limits (including systematic uncertainties) on WIMPs annihilating to muons, using dwarf spheroidal galaxies as a target and assuming the NFW profile. The limits shown here are from the Fermi-LAT combined analysis of 15 dwarf galaxies [124], observations of Segue 1 by MAGIC [129], and the IceCube combined analysis of 5 dwarf galaxies (Segue 1, Ursa Major II, Willman 1, Coma Berenices and Draco) described in this thesis.

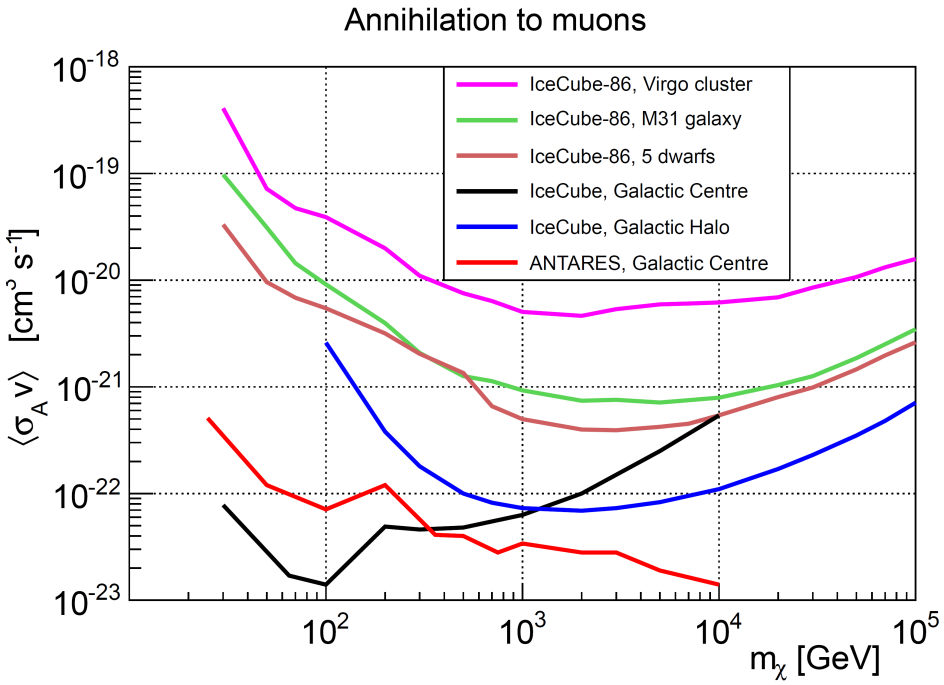


Figure 6.12.: Comparison of limits (including systematic uncertainties) on WIMPs annihilating to muons from neutrino experiments (IceCube [120,122] and ANTARES [121]), assuming an NFW profile without substructure. The limits from the analysis described in this thesis are denoted by 'IceCube-86'.

7. Summary and outlook

The goal of the research described in this thesis was to search for neutrinos originating from dark matter annihilations in nearby galaxies and galaxy clusters: the Virgo galaxy cluster, the M31 galaxy and five dwarf galaxies. If an excess of neutrinos from the direction of these targets were to be found, this would be evidence that dark matter exists and consists of particles, which are both topics that are still up for debate. However, no statistically significant excess was found, which resulted in limits being placed on the velocity-averaged annihilation cross section of dark matter particles to several annihilation channels.

For this analysis, three years of data from the completed IceCube detector were used. This dataset contained a large number of background events, both from atmospheric muons (reducible background) and atmospheric neutrinos (irreducible background), and possibly a number of signal events (neutrinos resulting from dark matter annihilations). As a first step, an event selection was applied which removed most of the atmospheric muons while keeping as much of the signal as possible, and this event selection resulted in a dataset dominated by neutrino events. An unbinned maximum likelihood analysis was applied to this dataset to determine whether there was an excess coming from the direction of the targets that were studied. This maximum likelihood analysis was done under the assumption that the target would be a point source. For the M31 galaxy and the Virgo cluster, an extended (Gaussian) profile was also studied.

An excess was found for three of the five dwarf galaxies, however the probability of an excess of this size or larger appearing in the data was 4.9%, so this was not statistically significant. As a result, limits were placed on the velocity-averaged dark matter annihilation cross section of dark matter particles with masses between 30 GeV and 100 TeV for five annihilation channels. For annihilation to muons, these limits reached down to $3.4 \cdot 10^{-22}$ cm^3 s^{-1} for WIMP masses of 2 and 3 TeV. Also, no evidence was found for an extended neutrino emission from the M31 galaxy or the Virgo cluster.

The resulting limits on the velocity-averaged dark matter annihilation cross section were improved by up to an order of magnitude compared to the previous analysis using similar targets. There were several reasons for this improvement, some of which were improvements in the analysis itself, while others were general improvements to the IceCube detector. The first improvement to this analysis specifically was that the event selection was done specifically for this analysis and used improved methods (some of which were used for the first time), for example taking into account a number of new event variables with clearly different distributions for signal and background. Also, the analysis method itself was improved significantly: instead of a binned analysis, an unbinned maximum likelihood method was used, which takes into account the expected error of the reconstruction and the reconstructed energy of the event. This analysis is also the first IceCube analysis where the Probability Density Function for the expected error of the reconstruction σ was included in the likelihood.

There were also general improvements to the IceCube detector: this analysis used three years of data instead of the one year of data that was used for the previous analysis. Also, the detector that was used for this analysis was larger than the detector used for the previous analysis (86 strings instead of 59 strings), and contained the low-energy subdetector DeepCore. This subdetector made it possible to set limits on the annihilation cross section for WIMP masses down to 30 GeV instead of 300 GeV. Third, the direction reconstruction used for this analysis performs significantly better than that used in the previous analysis.

In the coming years, the result is expected to be improved even further: the IceCube detector is still taking data, which means that the amount of data included in the analysis can be increased. A significant improvement would also be expected from extending the IceCube detector, which is currently under discussion [221]. Similarly, the result at low WIMP masses would be improved by the addition of PINGU, a planned very-low-energy subdetector [222]. In addition, research into direction and energy reconstruction methods is still ongoing, and especially improvements in the quality of the direction reconstruction would improve the results of this analysis. One interesting possible improvement to the analysis method would also be to add the output of the final BDT to the likelihood (as an extra term) instead of using a cut on the BDT output.

As has been mentioned in chapter 2, the search for additional dwarf galaxies is ongoing, and there are a number of candidates which could potentially be used in a future analysis. Combining data from observations of additional dwarf galaxies is another way to improve the results of this analysis, especially if (some of) these new dwarf galaxies would have high J-factors. Finally, as has been mentioned before, there are still uncertainties in the determination of the J-factors of the dwarf galaxies and other sources. Improving on this might not improve the limits on the velocity-averaged dark matter annihilation cross section, but it would improve the amount of trust that can be put in those limits.

The research described in this thesis is part of a large-scale search for dark matter, which has been going on for several decades and consists of searches using many different methods. So far, a number of possible signals have been found, but none have been confirmed. However, there are many promising searches still going on at the moment, so it is a very exciting time to be involved in dark matter research.

A. Event variables, IceCube-dominated sample, after preliminary cuts

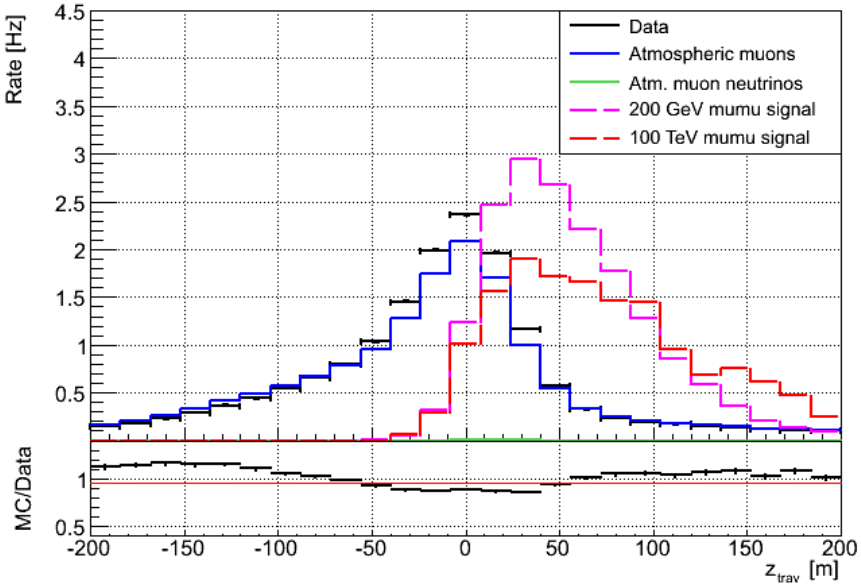


Figure A.1.: Histogram of z_{trav}, the average drift in the z direction with respect to the first quarter of hits, for the IceCube-dominated subsample after the preliminary cuts. Shown are a subset of the experimental data, simulated atmospheric muons and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

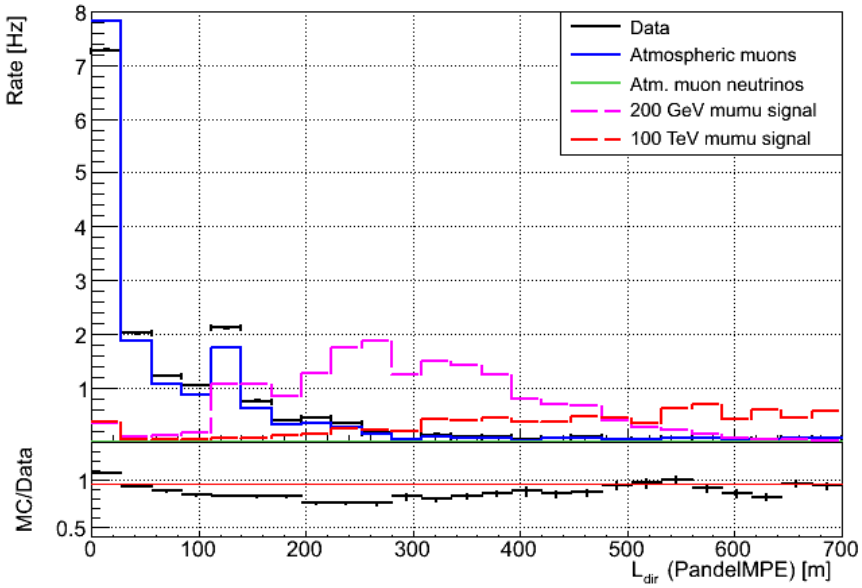


Figure A.2.: Histogram of L_{dir}, the track length along which there are direct hits (for the PandelMPE reconstruction), for the IceCube-dominated subsample after the preliminary cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

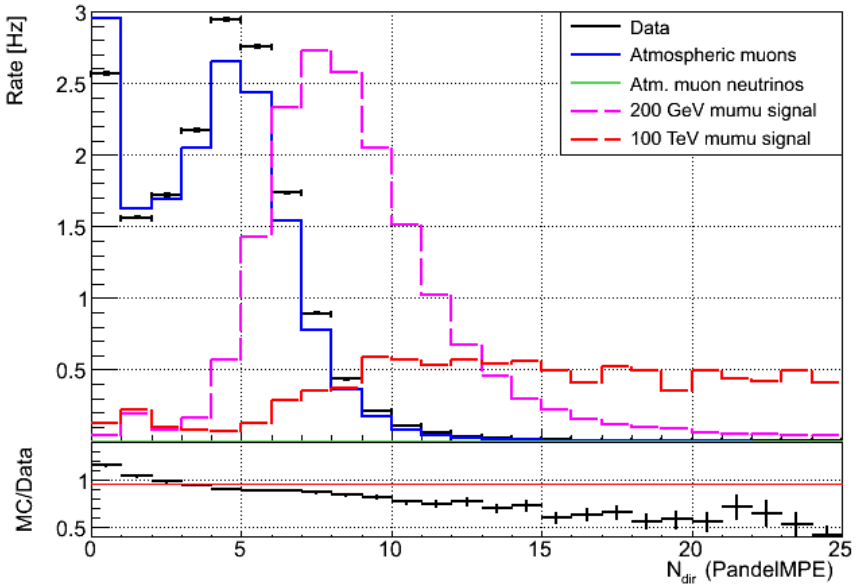


Figure A.3.: Histogram of N_{dir}, the number of direct hits (for the PandelMPE reconstruction), for the IceCube-dominated subsample after the preliminary cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

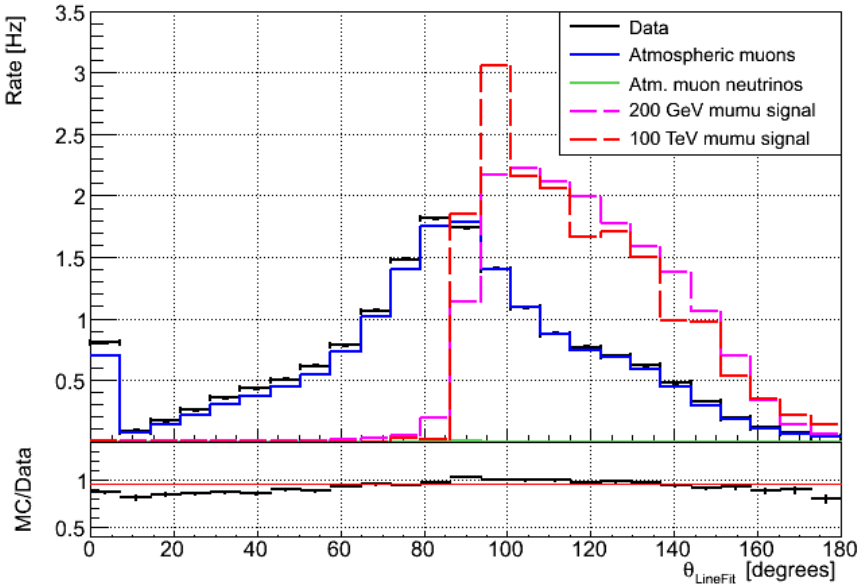


Figure A.4.: Histogram of θ_{LF}, the zenith as reconstructed by LineFit, for the IceCube-dominated subsample after the preliminary cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

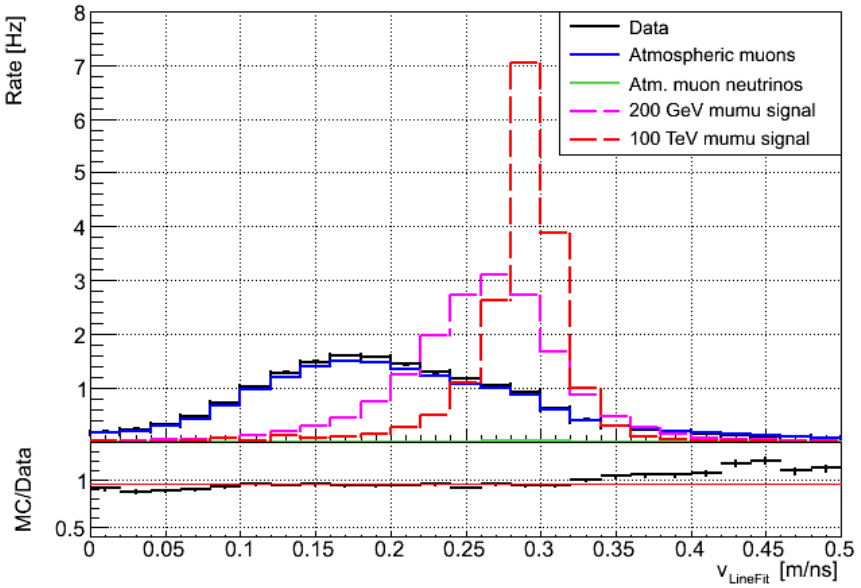


Figure A.5.: Histogram of v_{LF}, the speed as reconstructed by LineFit, for the IceCube-dominated subsample after the preliminary cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

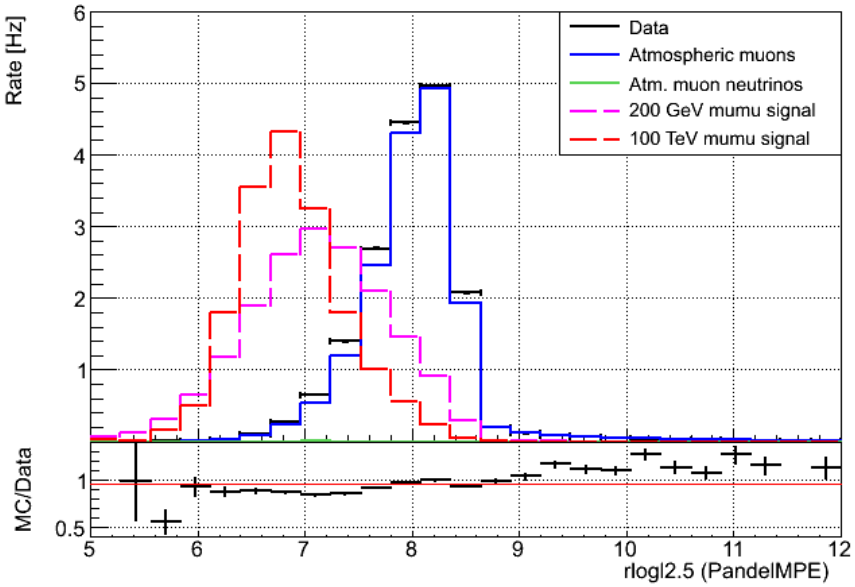


Figure A.6.: Histogram of rlogl2.5, the reduced log-likelihood (for the PandelMPE reconstruction) for the IceCube-dominated subsample after the preliminary cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

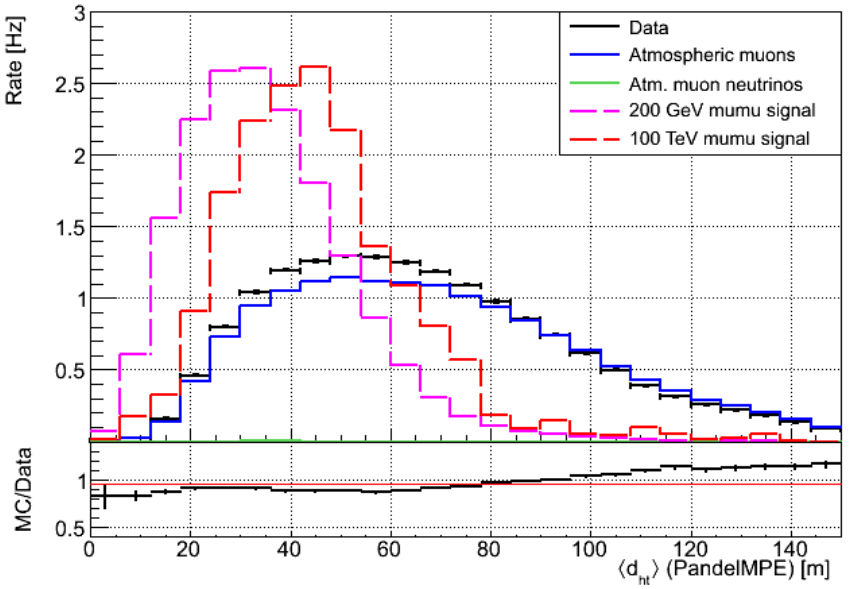


Figure A.7.: Histogram of $\langle d_{ht} \rangle$, the charge-weighted average track-DOM distance (for the PandelMPE reconstruction), for the IceCube-dominated subsample after the preliminary cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

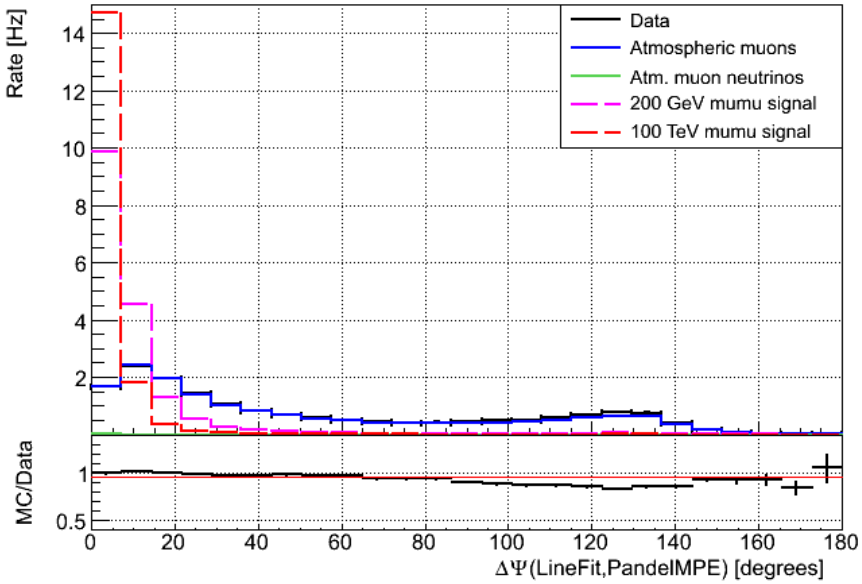


Figure A.8.: Histogram of $\Delta\Psi$(LineFit,PandelMPE), the angle between the tracks as reconstructed by LineFit and PandelMPE, for the IceCube-dominated subsample after the preliminary cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

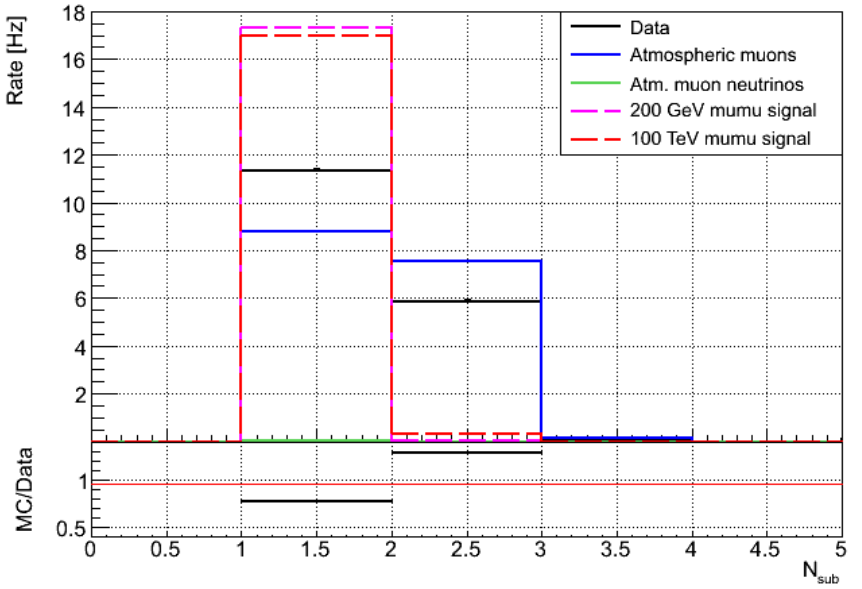


Figure A.9.: Histogram of N_{sub}, the number of subevents found by the topological splitting algorithm, for the IceCube-dominated subsample after the preliminary cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

B. Event variables, DeepCore-dominated sample, after second level of cuts

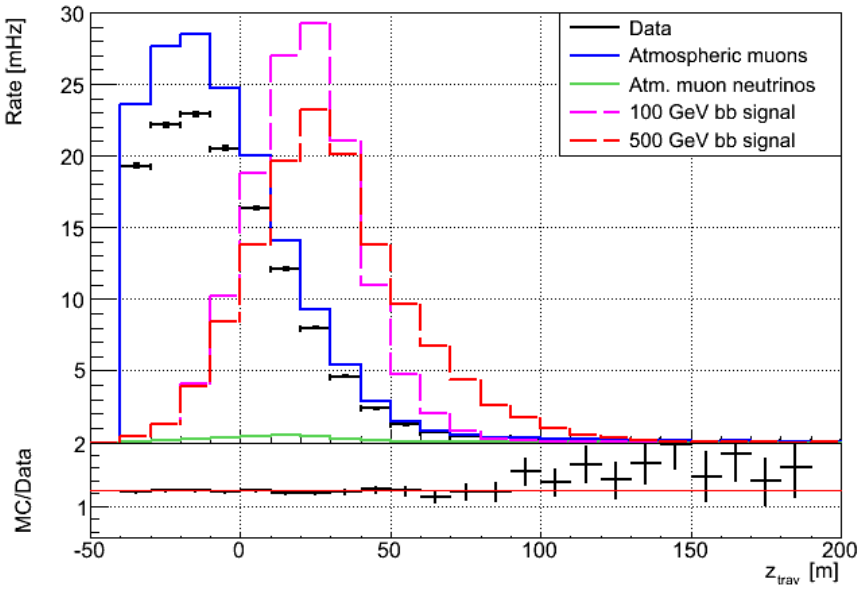


Figure B.1.: Histogram of z_{trav}, the average drift in the z direction with respect to the first quarter of hits, for the DeepCore-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs and 100 GeV WIMPs, both annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

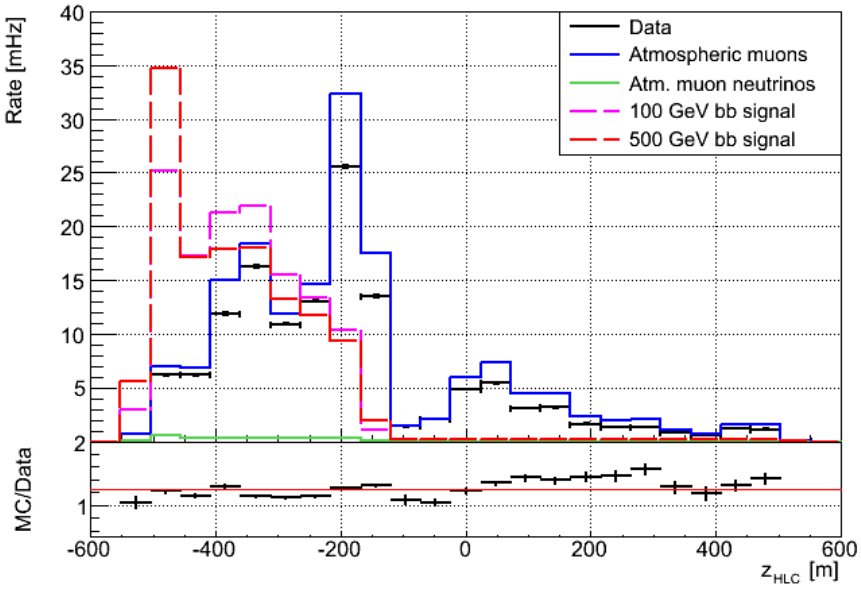


Figure B.2.: Histogram of z_{HLC}, the z position of the first HLC hit, for the DeepCore-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs and 100 GeV WIMPs, both annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

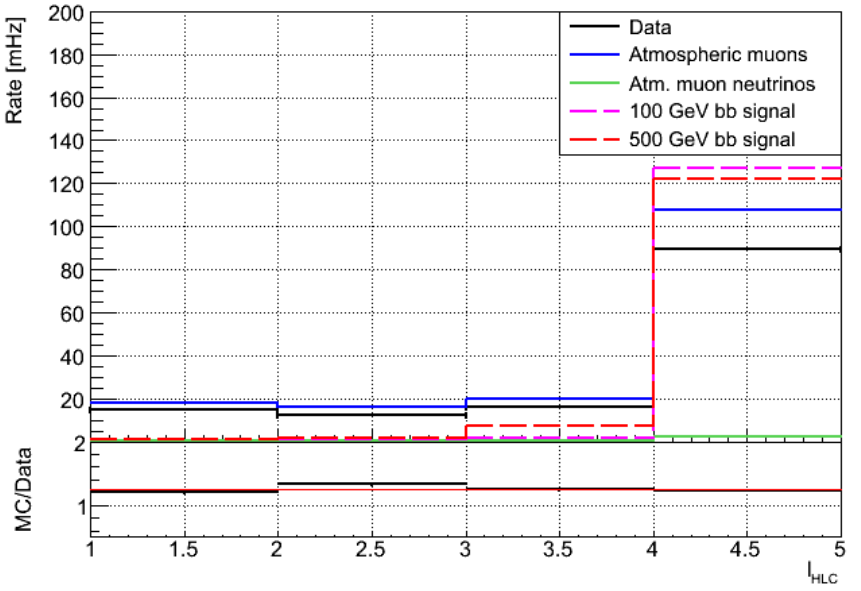


Figure B.3.: Histogram of l_{HLC}, the string layer that the DOM that detected the first HLC hit is on, for the DeepCore-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs and 100 GeV WIMPs, both annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

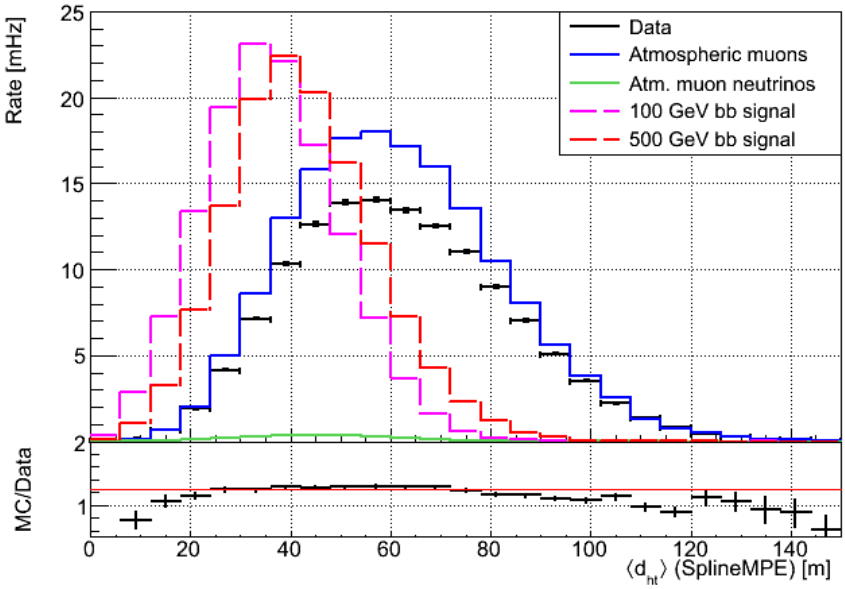


Figure B.4.: Histogram of $\langle d_{ht} \rangle$, the charge-weighted average track-DOM distance (for the SplineMPE reconstruction), for the DeepCore-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs and 100 GeV WIMPs, both annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

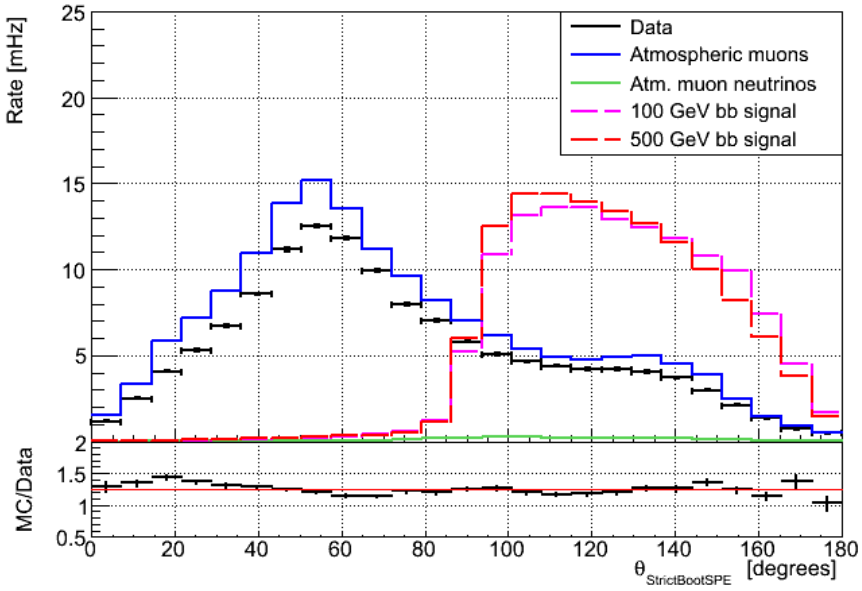


Figure B.5.: Histogram of $\theta_{StrictBootSPE}$, the zenith as reconstructed by BootSPE, using a stricter hit cleaning than usual, for the DeepCore-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs and 100 GeV WIMPs, both annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

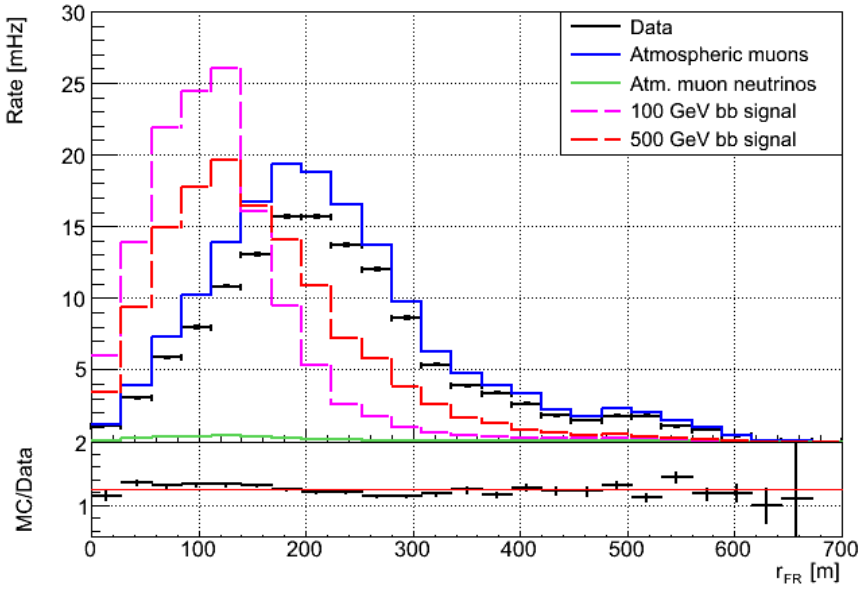


Figure B.6.: Histogram of r_{FR}, the r value of the track starting point as reconstructed by FiniteReco, for the DeepCore-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs and 100 GeV WIMPs, both annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

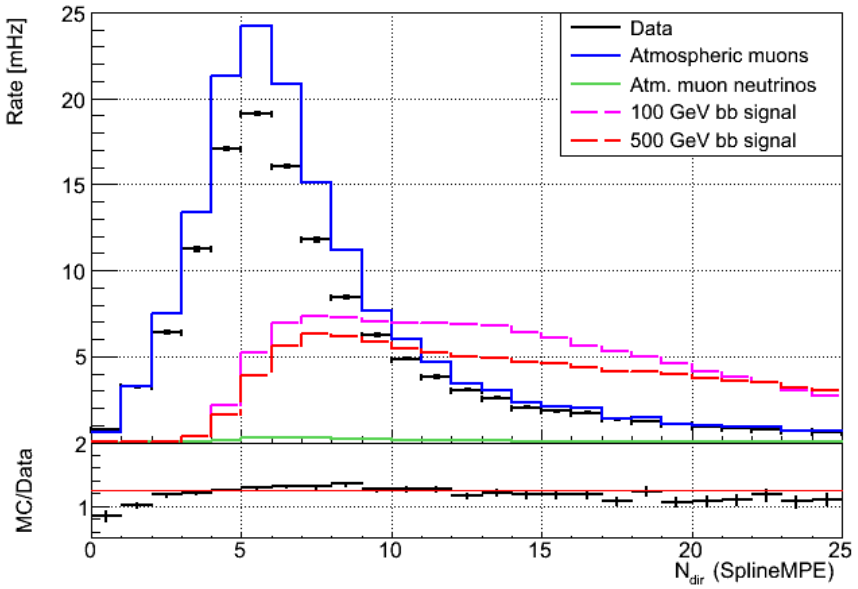


Figure B.7.: Histogram of N_{dir}, the number of direct hits (for the SplineMPE reconstruction), for the DeepCore-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs and 100 GeV WIMPs, both annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

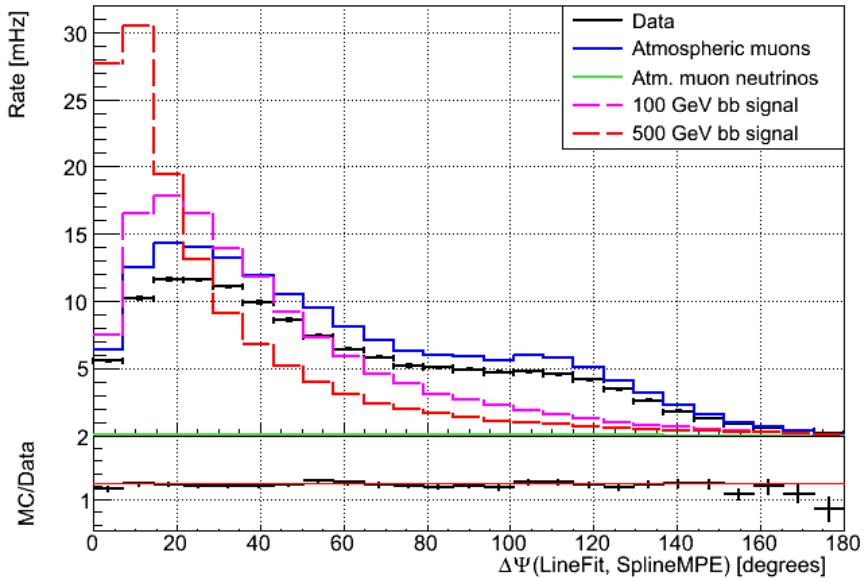


Figure B.8.: Histogram of $\Delta\Psi$(LineFit, SplineMPE), the angle between the tracks as reconstructed by LineFit and SplineMPE, for the DeepCore-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs and 100 GeV WIMPs, both annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

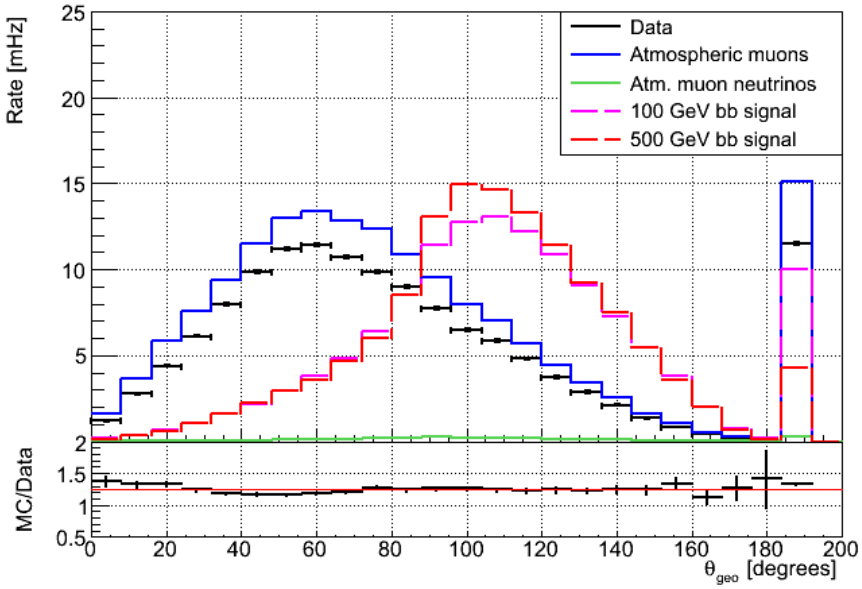


Figure B.9.: Histogram of θ_{geo}, the smallest zenith (as reconstructed by PandelSPE) of the geometrically split subevents, for the DeepCore-dominated subsample after the second level of cuts. In some cases, a subevent does not contain enough hits to be reconstructed by PandelSPE, in that case its zenith is artificially set to 190 degrees. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (500 GeV WIMPs and 100 GeV WIMPs, both annihilating to b quarks) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

C. Event variables, IceCube-dominated sample, after second level of cuts

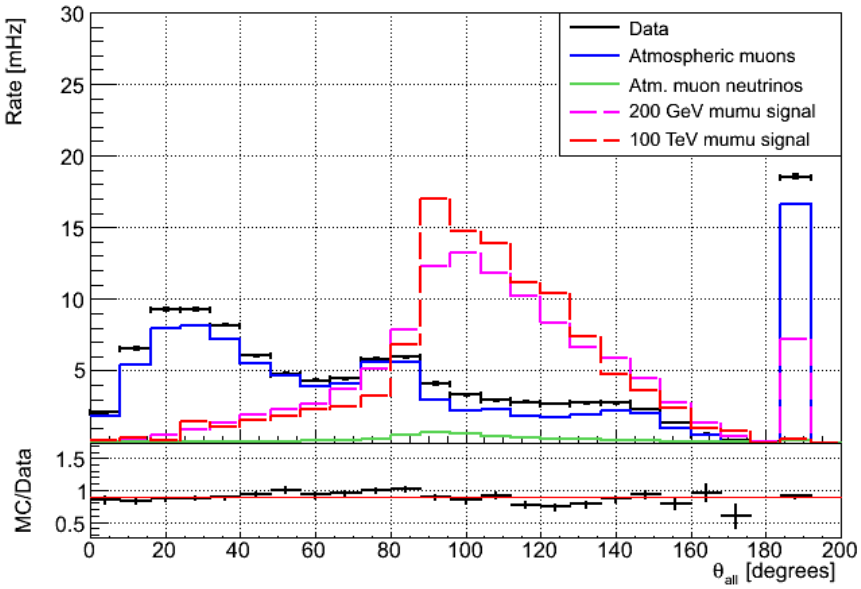


Figure C.1.: Histogram of θ_{all}, the smallest zenith (as reconstructed by PandelSPE) of all subevents (both geometrically split and split in time), for the IceCube-dominated subsample after the second level of cuts. In some cases, a subevent does not contain enough hits to be reconstructed by PandelSPE, in that case its zenith is artificially set to 190 degrees. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

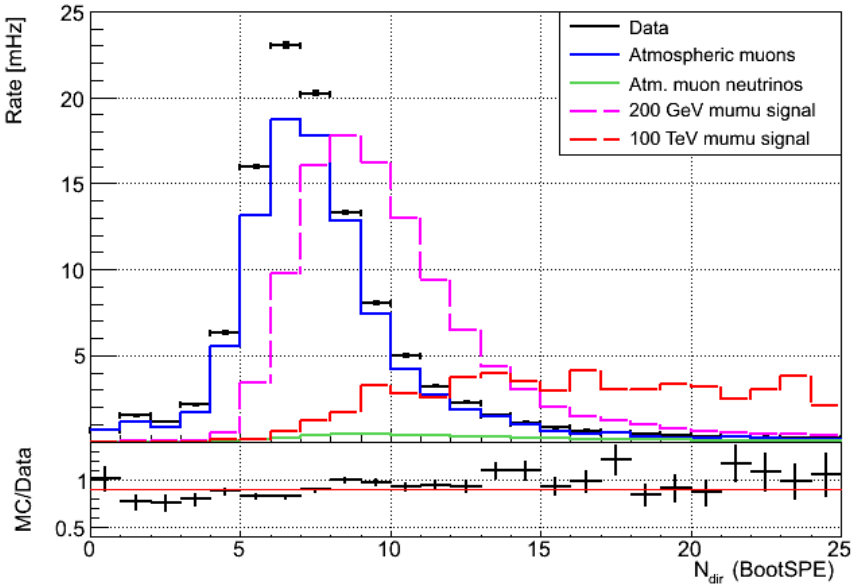


Figure C.2.: Histogram of N_{dir}, the number of direct hits (for the BootSPE reconstruction), for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

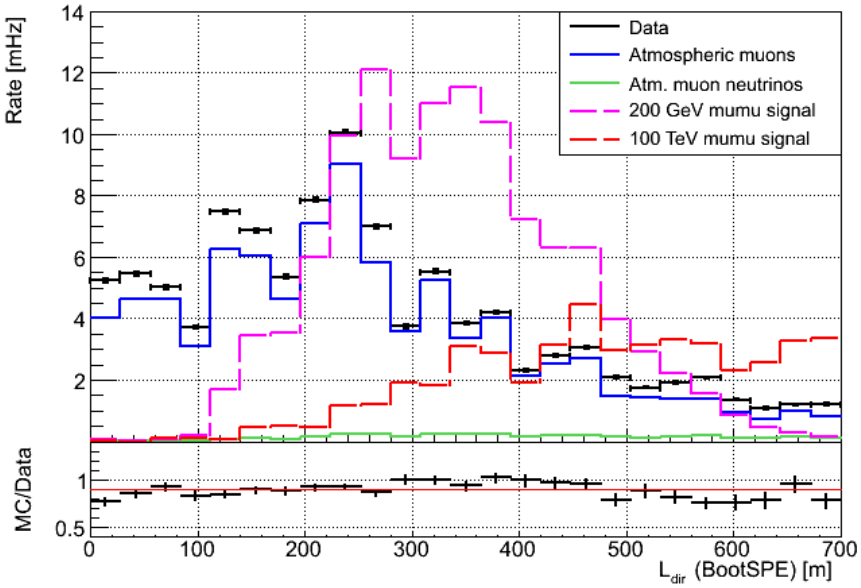


Figure C.3.: Histogram of L_{dir}, the track length along which there are direct hits (for the BootSPE reconstruction), for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

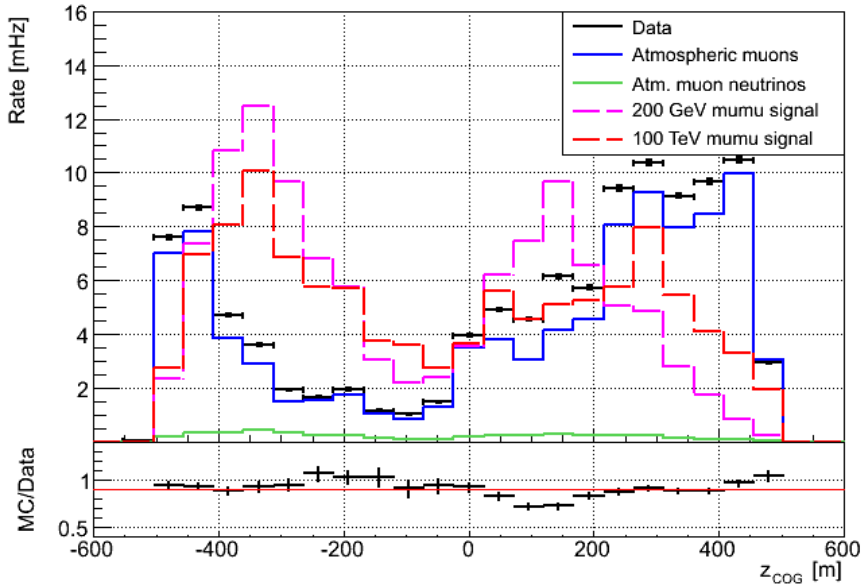


Figure C.4.: Histogram of z_{COG}, the z position of the Center of Gravity, for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

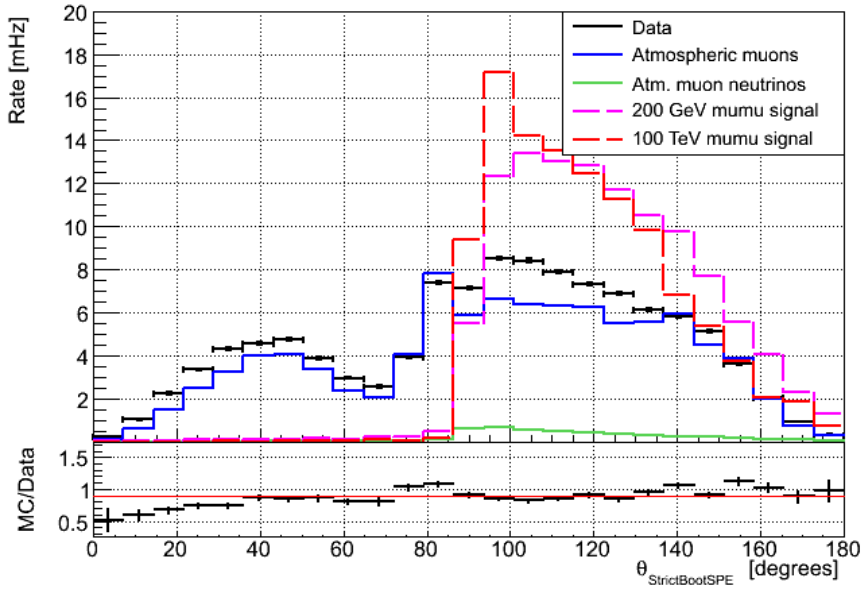


Figure C.5.: Histogram of $\theta_{StrictBootSPE}$, the zenith as reconstructed by BootSPE, using a stricter hit cleaning, for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

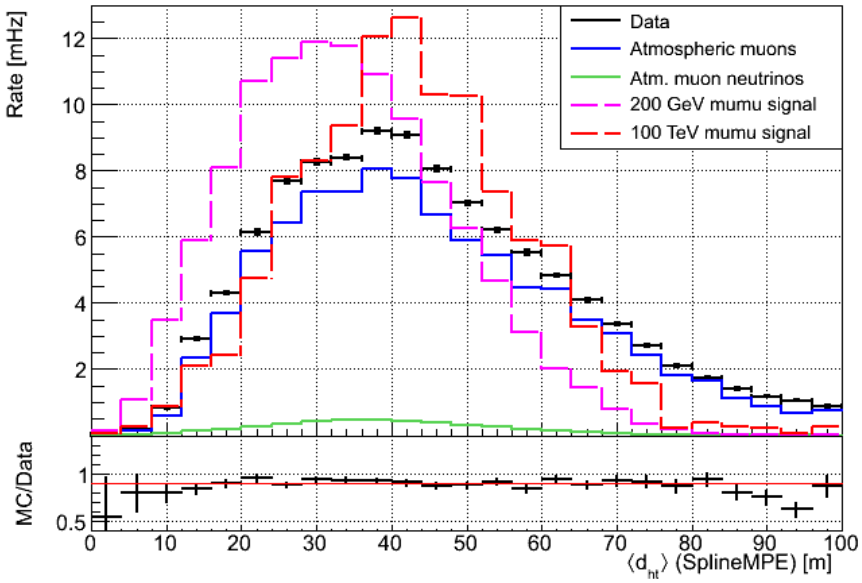


Figure C.6.: Histogram of $\langle d_{ht} \rangle$, the charge-weighted average track-to-DOM distance (for the SplineMPE reconstruction), for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

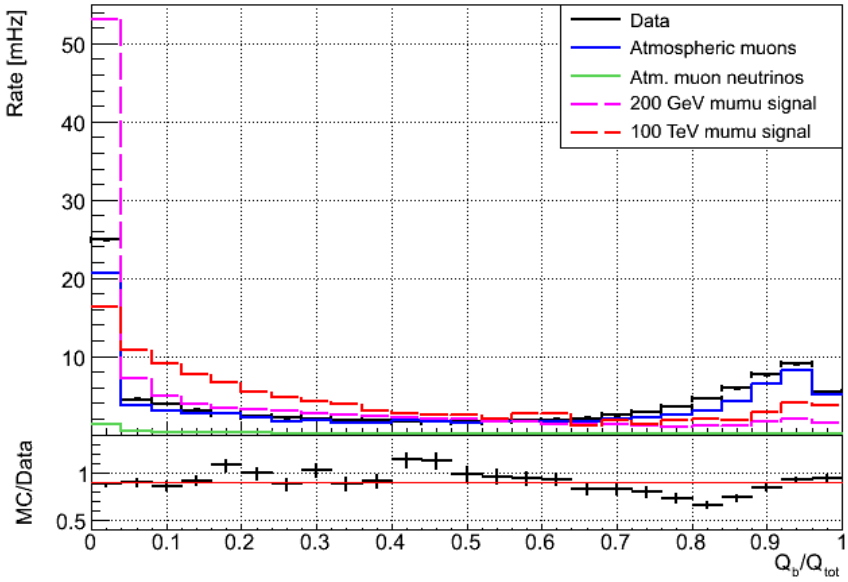


Figure C.7.: Histogram of Q_b/Q_{tot}, the fraction of the total charge that is on the boundary layer of the detector, for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

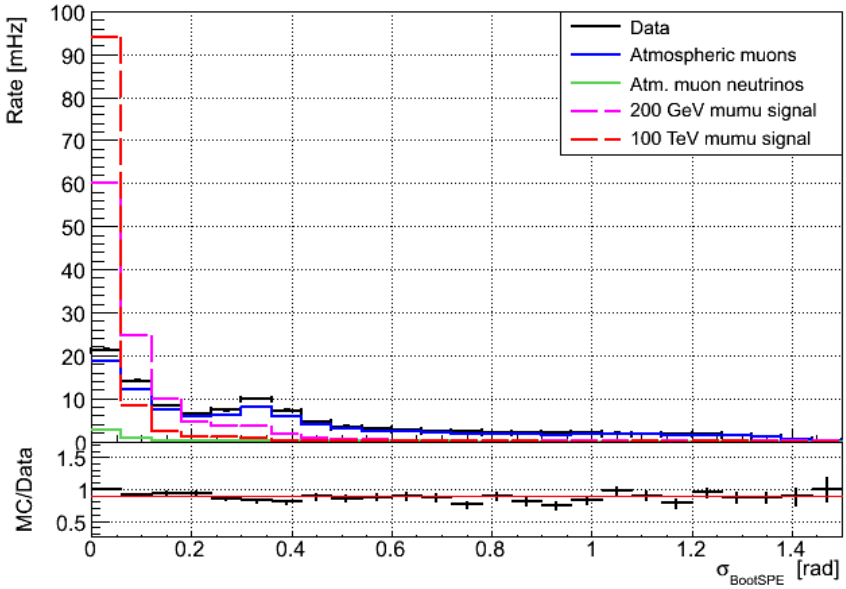


Figure C.8.: Histogram of $\sigma_{BootSPE}$, the estimated error of the BootSPE reconstruction, for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

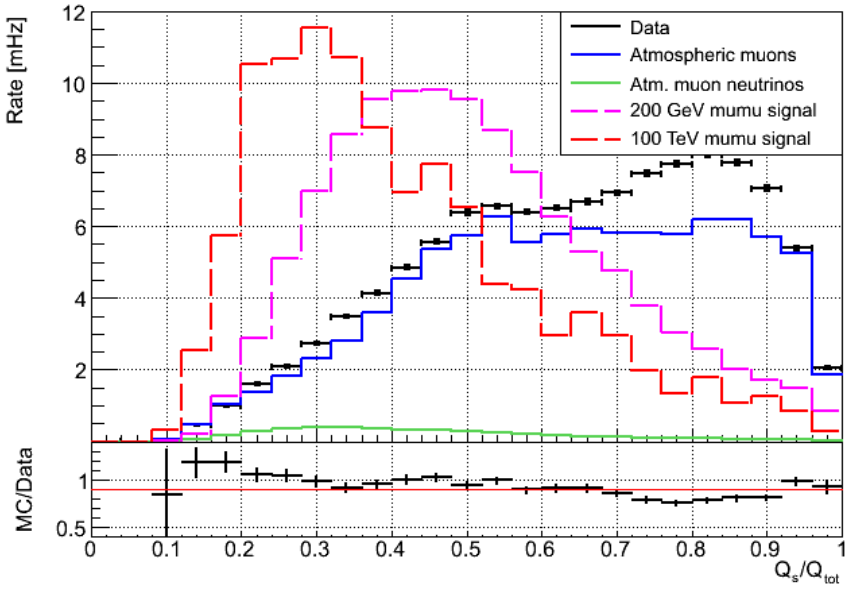


Figure C.9.: Histogram of Q_s/Q_{tot}, the fraction of the total charge that is on the string with the highest charge, for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

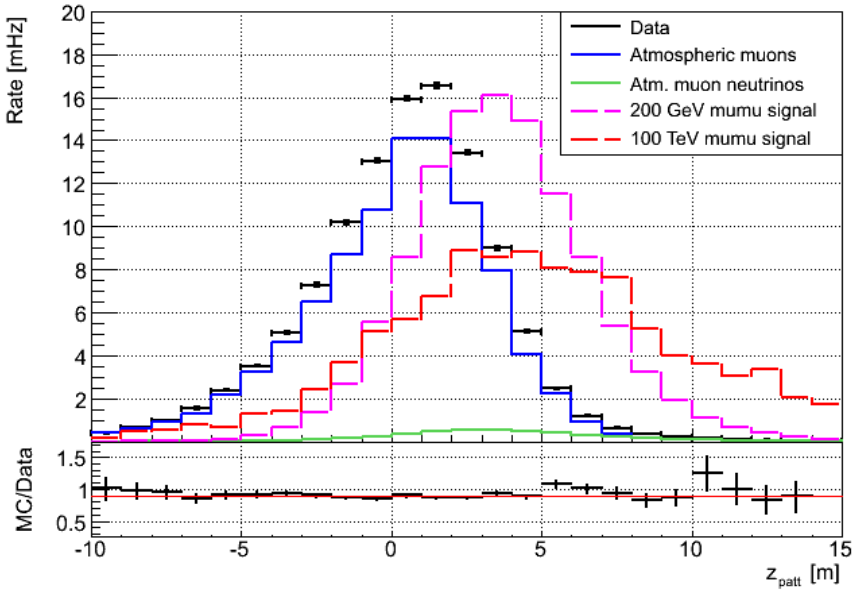


Figure C.10.: Histogram of z_{patt}, the z pattern, for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

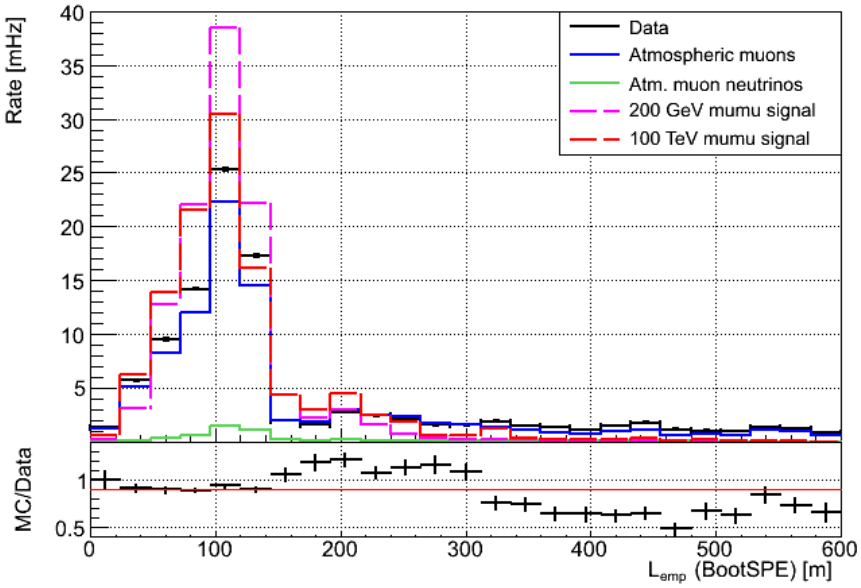


Figure C.11.: Histogram of L_{emp}, the maximum length along the track (as reconstructed by BootSPE) with no hit DOMs within a radius of 150 meters, for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

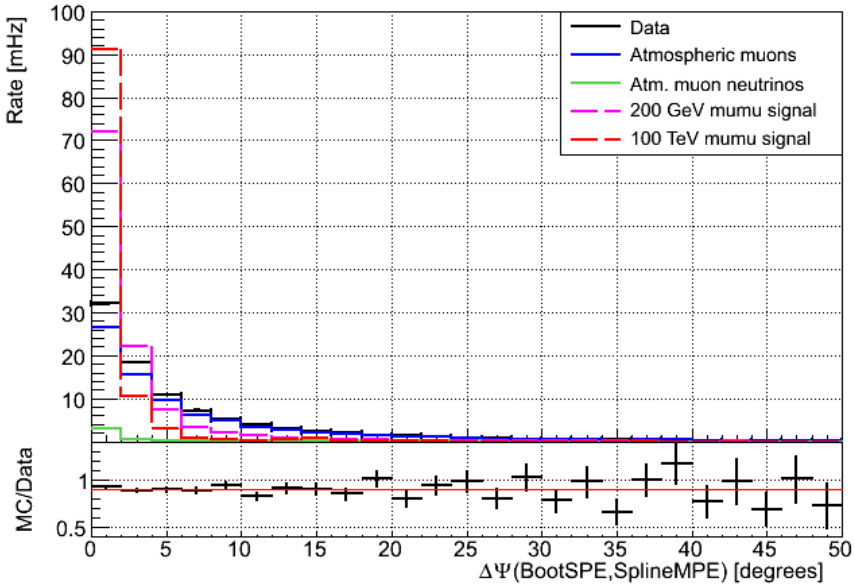


Figure C.12.: Histogram of $\Delta\Psi$(SplineMPE, BootSPE), the angle between the tracks as reconstructed by SplineMPE and BootSPE, for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

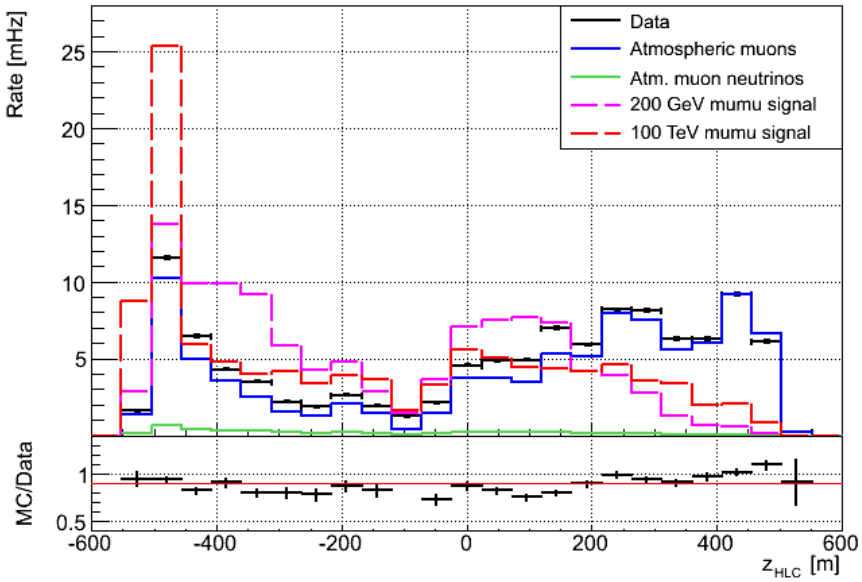


Figure C.13.: Histogram of z_{HLC}, the z position of the first HLC hit, for the IceCube-dominated subsample after the second level of cuts. Shown are a subset of the experimental data, simulated atmospheric muon and atmospheric muon neutrinos, and two signal samples (100 TeV WIMPs and 200 GeV WIMPs, both annihilating to muons) which are approximately normalized to the data rate. Below the main plot, the ratio of Monte Carlo simulation (the sum of the atmospheric muons and the atmospheric muon neutrinos) and data is shown.

D. Results for dwarf galaxies

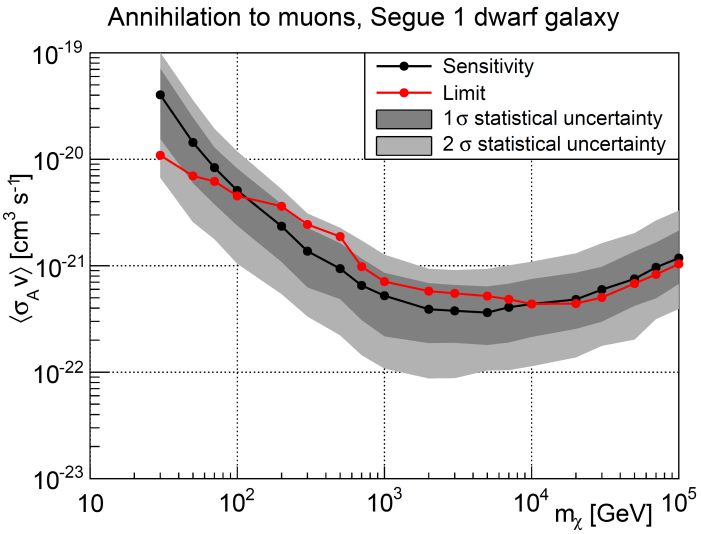


Figure D.1.: Comparing the sensitivity (black) and 90% C.L. limit (red) on the velocity-averaged WIMP annihilation cross section for annihilation to muons for the Segue 1 dwarf galaxy. The statistical spread on the sensitivity is also shown (grey one- and two-sigma bands). The markers denote which WIMP masses have been tested, the lines have been added to guide the eye. Systematic uncertainties have not yet been taken into account for these limits.

m_χ [GeV]	$\hat{\mu}$	$\mu_{l,90\%}$	$\mu_{h,90\%}$	$\langle\sigma_A v\rangle_{lim}$ [10^{-21} cm^3 s^{-1}]	$\langle\sigma_A v\rangle_{sens}$ [10^{-21} cm^3 s^{-1}]
30	0.0	0.0	2.6	11	40
50	0.0	0.0	6.8	7.0	14
70	0.0	0.0	11.5	6.2	8.4
100	0.0	0.0	16.0	4.5	5.1
200	14.9	0.0	40.0	3.6	2.4
300	18.1	3.0	41.0	2.4	1.4
500	17.6	4.8	38.0	1.9	0.94
700	4.1	0.0	21.0	0.98	0.65
1000	4.3	0.0	19.0	0.71	0.52
2000	4.2	0.0	18.5	0.58	0.39
3000	3.1	0.0	17.5	0.55	0.38
5000	0.66	0.0	15.0	0.52	0.36
7000	0.0	0.0	12.5	0.48	0.41
10000	0.0	0.0	9.6	0.44	0.44
20000	0.0	0.0	6.4	0.44	0.48
30000	0.0	0.0	5.4	0.50	0.60
50000	0.0	0.0	4.7	0.68	0.75
70000	0.0	0.0	4.2	0.83	0.97
100000	0.0	0.0	3.7	1.0	1.2

Table D.1.: The best fit and 90% C.L. lower and upper limits for the number of signal events μ in the total dataset, and the sensitivity to and 90% C.L. upper limit on the velocity-averaged annihilation cross section $\langle\sigma_A v\rangle$, calculated using the method described in sections 5.1 and 5.5, for annihilation to muons and the Segue dwarf galaxy, for different WIMP masses m_χ. Systematic uncertainties have not yet been taken into account for these limits.

m_χ [GeV]	$\hat{\mu}$	$\mu_{l,90\%}$	$\mu_{h,90\%}$	$\langle\sigma_A v\rangle_{lim}$ [10^{-21} cm^3 s^{-1}]	$\langle\sigma_A v\rangle_{sens}$ [10^{-21} cm^3 s^{-1}]
30	0.0	0.0	9.6	42	55
50	0.22	0.0	13.5	17	19
70	3.1	0.0	17.0	12	11
100	6.5	0.0	27.0	10	6.5
200	25.1	6.6	56.0	7.1	2.9
300	23.9	7.4	54.0	4.5	1.6
500	19.6	6.0	44.0	2.8	1.1
700	3.7	0.0	19.0	0.97	0.79
1000	4.0	0.0	18.0	0.79	0.59
2000	3.9	0.0	17.0	0.70	0.49
3000	4.2	0.0	17.0	0.74	0.48
5000	4.7	0.0	17.0	0.87	0.49
7000	4.6	0.0	15.5	0.92	0.52
10000	4.6	0.6	15.0	1.1	0.58
20000	3.6	0.9	12.0	1.5	0.76
30000	3.0	0.7	10.5	1.9	0.95
50000	2.4	0.6	9.0	2.8	1.4
70000	2.2	0.6	8.4	3.9	1.9
100000	2.0	0.6	7.4	5.4	2.7

Table D.2.: The best fit and 90% C.L. lower and upper limits for the number of signal events μ in the total dataset, and the sensitivity to and 90% C.L. upper limit on the velocity-averaged annihilation cross section $\langle\sigma_A v\rangle$, calculated using the method described in sections 5.1 and 5.5, for annihilation to muons and the Ursa Major II dwarf galaxy, for different WIMP masses m_χ. Systematic uncertainties have not yet been taken into account for these limits.

m_χ [GeV]	$\hat{\mu}$	$\mu_{l,90\%}$	$\mu_{h,90\%}$	$\langle\sigma_A v\rangle_{lim}$ [10^{-21} cm^3 s^{-1}]	$\langle\sigma_A v\rangle_{sens}$ [10^{-21} cm^3 s^{-1}]
30	6.4	0.5	17.5	$1.4 \cdot 10^2$	75
50	0.97	0.0	15.0	30	21
70	0.0	0.0	12.0	12	12
100	0.0	0.0	14.0	7.3	9.2
200	0.0	0.0	5.2	0.94	3.8
300	0.0	0.0	2.5	0.30	2.5
500	0.0	0.0	2.0	0.19	1.5
700	0.0	0.0	15.5	0.98	1.2
1000	0.0	0.0	16.0	0.91	0.91
2000	1.8	0.0	17.0	0.95	0.73
3000	3.8	0.0	18.0	1.1	0.73
5000	5.3	0.0	18.5	1.3	0.76
7000	5.5	0.2	18.5	1.6	0.79
10000	5.2	0.5	17.5	1.8	0.82
20000	4.2	0.4	14.5	2.5	1.2
30000	3.5	0.5	12.5	3.0	1.4
50000	2.7	0.4	11.0	4.4	1.9
70000	2.4	0.3	9.8	5.7	2.9
100000	2.1	0.3	9.2	8.2	3.9

Table D.3.: The best fit and 90% C.L. lower and upper limits for the number of signal events μ in the total dataset, and the sensitivity to and 90% C.L. upper limit on the velocity-averaged annihilation cross section $\langle\sigma_A v\rangle$, calculated using the method described in sections 5.1 and 5.5, for annihilation to muons and the Willman 1 dwarf galaxy, for different WIMP masses m_χ. Systematic uncertainties have not yet been taken into account for these limits.

m_χ [GeV]	$\hat{\mu}$	$\mu_{l,90\%}$	$\mu_{h,90\%}$	$\langle\sigma_A v\rangle_{lim}$ [10^{-21} cm^3 s^{-1}]	$\langle\sigma_A v\rangle_{sens}$ [10^{-21} cm^3 s^{-1}]
30	2.4	0.0	13.0	$1.4 \cdot 10^2$	$1.0 \cdot 10^2$
50	3.9	0.0	18.5	52	36
70	7.1	0.0	23.0	35	20
100	14.4	2.9	34.0	30	14
200	13.0	0.0	37.0	13	6.9
300	6.4	0.0	25.0	6.1	4.3
500	2.7	0.0	18.0	3.8	3.2
700	0.0	0.0	3.4	0.77	2.5
1000	0.0	0.0	2.9	0.51	1.7
2000	0.0	0.0	2.2	0.30	1.3
3000	0.0	0.0	2.3	0.30	1.2
5000	0.0	0.0	2.6	0.36	1.2
7000	0.0	0.0	4.0	0.61	1.3
10000	0.0	0.0	6.4	1.1	1.3
20000	2.6	0.2	9.2	2.4	1.6
30000	3.2	0.7	9.4	3.2	1.9
50000	3.6	1.1	9.6	5.1	2.4
70000	3.7	1.1	9.6	7.0	3.2
100000	3.8	1.2	9.4	9.8	4.1

Table D.4.: The best fit and 90% C.L. lower and upper limits for the number of signal events μ in the total dataset, and the sensitivity to and 90% C.L. upper limit on the velocity-averaged annihilation cross section $\langle\sigma_A v\rangle$, calculated using the method described in sections 5.1 and 5.5, for annihilation to muons and the Coma Berenices dwarf galaxy, for different WIMP masses m_χ. Systematic uncertainties have not yet been taken into account for these limits.

m_χ [GeV]	$\hat{\mu}$	$\mu_{l,90\%}$	$\mu_{h,90\%}$	$\langle\sigma_A v\rangle_{lim}$ [10^{-21} cm^3 s^{-1}]	$\langle\sigma_A v\rangle_{sens}$ [10^{-21} cm^3 s^{-1}]
30	1.0	0.0	11.0	$1.7 \cdot 10^2$	$1.7 \cdot 10^2$
50	0.29	0.0	11.0	43	45
70	2.1	0.0	16.0	34	29
100	3.0	0.0	22.0	22	20
200	2.1	0.0	29.0	9.1	8.8
300	0.16	0.0	23.0	4.7	4.7
500	0.0	0.0	18.0	2.8	3.2
700	0.0	0.0	10.5	1.7	2.4
1000	0.0	0.0	7.8	1.1	1.7
2000	0.0	0.0	6.8	0.89	1.4
3000	0.0	0.0	6.8	0.94	1.4
5000	0.0	0.0	6.2	1.0	1.5
7000	0.0	0.0	6.4	1.2	1.6
10000	0.0	0.0	5.6	1.3	1.8
20000	0.0	0.0	4.8	1.8	2.1
30000	0.0	0.0	3.6	1.9	2.8
50000	0.0	0.0	3.1	2.7	4.1
70000	0.0	0.0	3.0	3.9	5.1
100000	0.0	0.0	2.8	5.7	8.0

Table D.5.: The best fit and 90% C.L. lower and upper limits for the number of signal events μ in the total dataset, and the sensitivity to and 90% C.L. upper limit on the velocity-averaged annihilation cross section $\langle\sigma_A v\rangle$, calculated using the method described in sections 5.1 and 5.5, for annihilation to muons and the Draco dwarf galaxy, for different WIMP masses m_χ. Systematic uncertainties have not yet been taken into account for these limits.

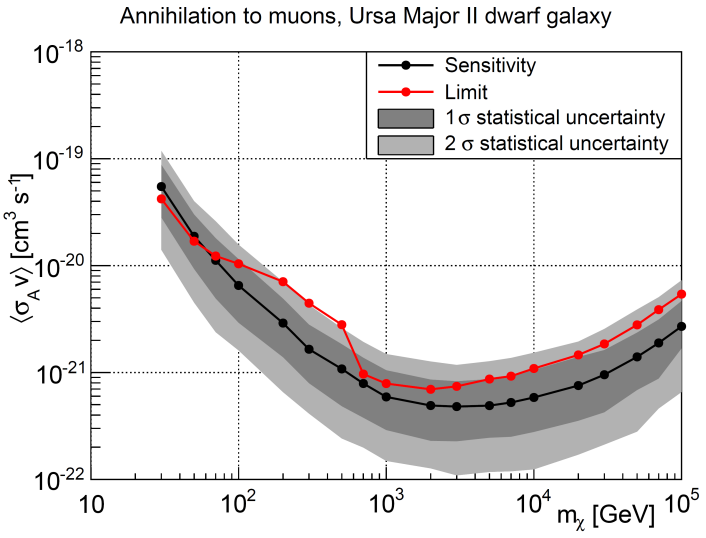


Figure D.2.: Comparing the sensitivity (black) and 90% C.L. limit (red) on the velocity-averaged WIMP annihilation cross section for annihilation to muons for the Ursa Major II dwarf galaxy. The statistical spread on the sensitivity is also shown (grey one- and two-sigma bands). The markers denote which WIMP masses have been tested, the lines have been added to guide the eye. Systematic uncertainties have not yet been taken into account for these limits.

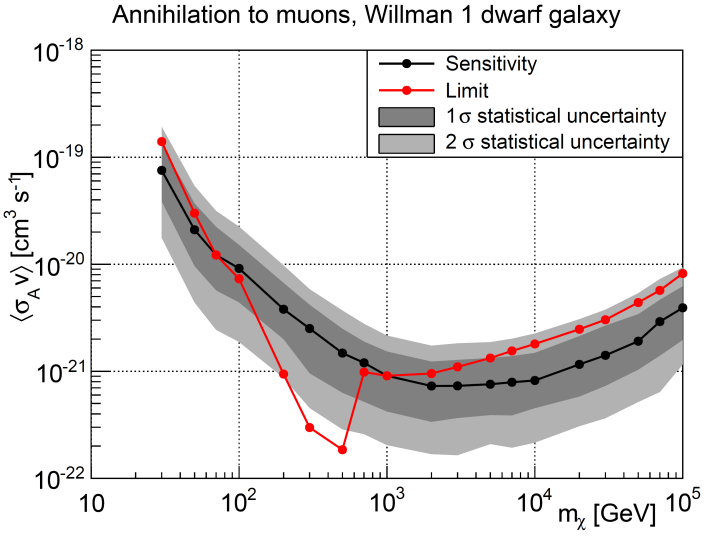


Figure D.3.: Comparing the sensitivity (black) and 90% C.L. limit (red) on the velocity-averaged WIMP annihilation cross section for annihilation to muons for the Willman 1 dwarf galaxy. The statistical spread on the sensitivity is also shown (grey one- and two-sigma bands). The markers denote which WIMP masses have been tested, the lines have been added to guide the eye. Systematic uncertainties have not yet been taken into account for these limits.

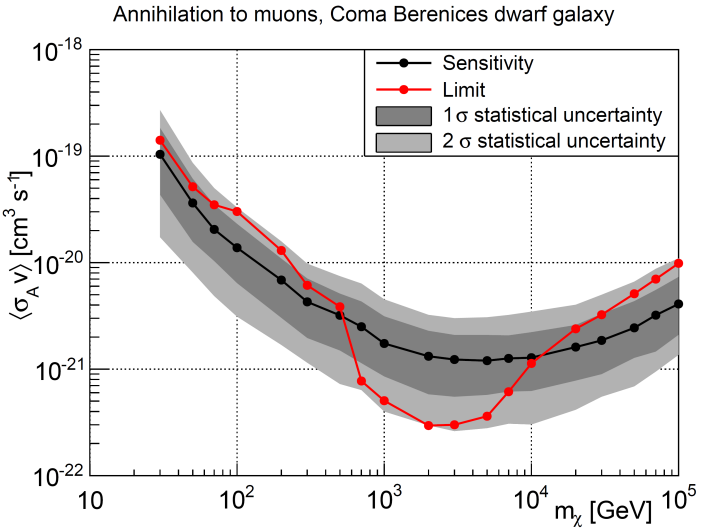


Figure D.4.: Comparing the sensitivity (black) and 90% C.L. limit (red) on the velocity-averaged WIMP annihilation cross section for annihilation to muons for the Coma Berenices dwarf galaxy. The statistical spread on the sensitivity is also shown (grey one- and two-sigma bands). The markers denote which WIMP masses have been tested, the lines have been added to guide the eye. Systematic uncertainties have not yet been taken into account for these limits.

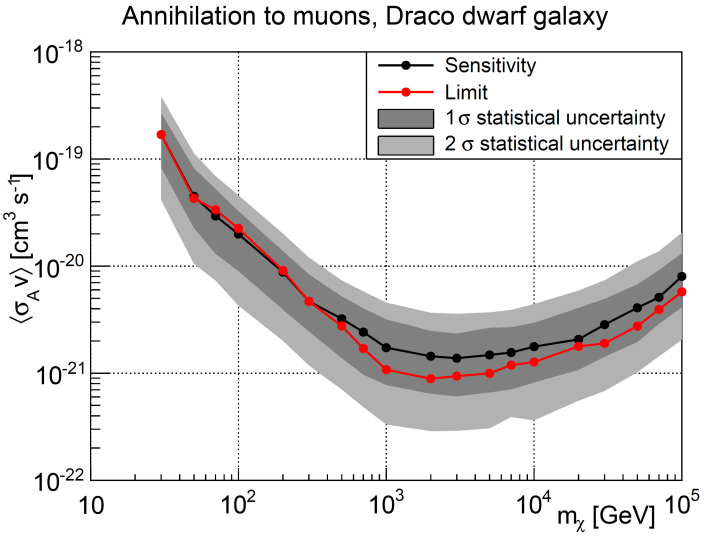


Figure D.5.: Comparing the sensitivity (black) and 90% C.L. limit (red) on the velocity-averaged WIMP annihilation cross section for annihilation to muons for the Draco dwarf galaxy. The statistical spread on the sensitivity is also shown (grey one- and two-sigma bands). The markers denote which WIMP masses have been tested, the lines have been added to guide the eye. Systematic uncertainties have not yet been taken into account for these limits.

Bibliography

[1] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, 1933.

[2] M. Milgrom. A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. *The Astrophysical Journal*, 270(365-370), 1983.

[3] G. Bertone, editor. *Particle Dark Matter*. Cambridge University Press, 2010.

[4] V.C. Rubin and W.K. Ford. Rotation of the Andromeda nebula from a spectroscopic survey of emission regions. *The Astrophysical Journal*, 159:379, 1970.

[5] K.G. Begeman, A.H. Broeils, and R.H. Sanders. Extended rotation curves of spiral galaxies: dark haloes and modified dynamics. *Monthly Notices of the Royal Astronomical Society*, 249:523–537, 1991.

[6] A. Einstein and F.A. Davis. *The Principle of Relativity*. Dover Books, 1952.

[7] W. Tucker et al. 1E 0657-56: a contender for the hottest known cluster of galaxies. *The Astrophysical Journal Letters*, 496(1):L5, 1998.

[8] D. Clowe et al. A direct empirical proof of the existence of dark matter. *The Astrophysical Journal*, 648:L109–L113, 2006.

[9] M. Markevitch et al. A textbook example of a bow shock in the merging galaxy cluster 1E 0657-56. *The Astrophysical Journal Letters*, 567(1):L27, 2002.

[10] B. Ryden. *Introduction to Cosmology*. Addison-Wesley, 2002.

[11] K.A. Olive et al. Review of particle physics. *Chinese Physics C*, 38(9):090001, 2014.

[12] C.L. Bennett et al. COBE Collaboration. Four-year COBE DMR Cosmic Microwave Background observations: maps and basic results. *The Astrophysical Journal Letters*, 464(1):L1, 1996.

[13] P.A.R. Ade et al. Planck Collaboration. Planck 2015 results. XIII. Cosmological parameters. 2015, arXiv:1502.01589.

[14] S. Das et al. ACT Collaboration. The Atacama Cosmology Telescope: temperature and gravitational lensing power spectrum measurements from three seasons of data. *Journal of Cosmology and Astroparticle Physics*, 2014(04):14, 2014.

[15] R. Keisler et al. SPT Collaboration. A measurement of the damping tail of the Cosmic Microwave Background power spectrum with the South Pole Telescope. *The Astrophysical Journal*, 743(1):28, 2011.

[16] D.H. Weinberg et al. Observational probes of cosmic acceleration. *Physics Reports*, 530(2):87–255, 2013.

[17] G. Bertone, D. Hooper, and J. Silk. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, 405:279–390, 2005.

[18] J.R. Bond and A.S. Szalay. The collisionless damping of density fluctuations in an expanding universe. *The Astrophysical Journal*, 274:443–468, 1983.

[19] P.J.E. Peebles. The origin of galaxies and clusters of galaxies. *Science*, 224(4656):1385–1391, 1984.

[20] C.A. Baker et al. Improved experimental limit on the electric dipole moment of the neutron. *Phys. Rev. Lett.*, 97:131801, 2006.

[21] R.D. Peccei and H.R. Quinn. CP conservation in the presence of pseudoparticles. *Physical Review Letters*, 38(25):1440, 1977.

[22] P. Sikivie. Experimental tests of the "invisible" axion. *Physical Review Letters*, 51:1415–1417, 1983.

[23] J. Hoskins et al. Search for nonvirialized axionic dark matter. *Physical Review D*, 84:121302, 2011.

[24] M. Arik et al. Search for solar axions by the CERN Axion Solar Telescope with He3 buffer gas: Closing the hot dark matter gap. *Phys. Rev. Lett.*, 112:091302, 2014.

[25] K. Ehret et al. New ALPS results on hidden-sector lightweights. *Physics Letters B*, 689:149–155, 2010.

[26] G. Steigman, B. Dasgupta, and J.F. Beacom. Precise relic WIMP abundance and its impact on searches for dark matter annihilation. *Physical Review D*, 86(2):023506, 2012.

[27] K. Griest and M. Kamionkowski. Unitarity limits on the mass and radius of dark-matter particles. *Phys. Rev. Lett.*, 64:615–618, 1990.

[28] E.W. Kolb, D.J.H Chung, and A. Riotto. WIMPZILLAS! In *DARK98: Proceedings of the Second International Conference on Dark Matter in Astro and Particle Physics*, 1998, arXiv:hep-ph/9810361.

[29] M. Drees, R.M. Godbole, and P. Roy. *Theory and phenomenology of sparticles.* World Scientific Publishing Company, 2005.

[30] H. Nishino et al. Super-Kamiokande Collaboration. Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in a large water Cherenkov detector. *Physical Review Letters*, 102:141801, 2009.

[31] K. Abe et al. Super-Kamiokande Collaboration. Search for proton decay via $p \to \nu k^+$ using 260 kiloton·year data of Super-Kamiokande. *Phys.Rev.D*, 90:072005, 2014.

[32] V. Takhistov et al. Super-Kamiokande Collaboration. Search for trilepton nucleon decay via $p \to e^+\nu\nu$ and $p \to \mu^+\nu\nu$ in the Super-Kamiokande experiment. *Phys. Rev. Lett.*, 113:101801, 2014.

[33] T. Falk, K.A. Olive, and M. Srednicki. Heavy sneutrinos as dark matter. *Phys.Lett.B*, 339:248–251, 1994.

[34] T. Kaluza. Zum Unitätsproblem in der Physik. *Sitzungsber. Preuss. Akad. Wiss. Berlin. (Math. Phys.)*, pages 966–972, 1921.

[35] O. Klein. Quantentheorie und fünfdimensionale Relativitätstheorie. *Zeitschrift für Physik*, 37:895–906, 1926.

[36] M.W. Goodman and E. Witten. Detectability of certain dark-matter candidates. *Physical Review D*, 31(12):3059, 1985.

[37] E. Aprile et al. XENON Collaboration. Dark matter results from 225 live days of XENON100 data. *Physical Review Letters*, 109(18):181301, 2012.

[38] D.S. Akerib et al. LUX Collaboration. First results from the LUX dark matter experiment at the Sanford Underground Research Facility. *Physical Review Letters*, 112(9):091303, 2014.

[39] R. Agnese et al. CDMS Collaboration. Search for low-mass weakly interacting massive particles with SuperCDMS. *Physical Review Letters*, 112(24):241302, 2014.

[40] R. Bernabei et al. DAMA Collaboration. Final model independent result of DAMA/LIBRA phase 1. *The European Physical Journal C*, 73(12):1–11, 2013.

[41] C.E. Aalseth et al. CoGeNT Collaboration. Results from a search for light-mass dark matter with a p-type point contact germanium detector. *Physical Review Letters*, 106(13):131301, 2011.

[42] G. Angloher et al. CRESST Collaboration. Results from 730 kg days of the CRESST-II dark matter search. *The European Physical Journal C*, 72(4):1–22, 2012.

[43] R. Agnese et al. CDMS Collaboration. Dark matter search results using the silicon detectors of CDMS II. *Physical Review Letters*, 2:013, 2013.

[44] G. Angloher et al. CRESST Collaboration. Results on low mass WIMPs using an upgraded CRESST-II detector. *The European Physical Journal C*, 74(12):1–6, 2014.

[45] E. Fernandez-Martinez and R. Mahbubani. The Gran Sasso muon puzzle. *Journal of Cosmology and Astroparticle Physics*, 2012(07):029, 2012.

[46] J.H. Davis. Fitting the annual modulation in DAMA with neutrons from muons and neutrinos. *Physical Review Letters*, 113(8):081302, 2014.

[47] A.L. Fitzpatrick, D. Hooper, and K.M. Zurek. Implications of CoGeNT and DAMA for light WIMP dark matter. *Phys.Rev.D*, 81:115005, 2010.

[48] J. Angle et al. XENON Collaboration. Search for light dark matter in XENON10 data. *Physical Review Letters*, 107(5):051301, 2011.

[49] Z. Ahmed et al. Combined limits on WIMPs from the CDMS and EDELWEISS experiments. *Physical Review D*, 84(1):011102, 2011.

[50] S.E.A. Orrigo. Direct dark matter search with XENON100. In *Proceedings of the 5th Roma International Conference on Astroparticle Physics RICAP-14*, arXiv:1501.03492, 2015.

[51] Z. Ahmed et al. CDMS Collaboration. Results from a low-energy analysis of the CDMS II germanium data. *Physical Review Letters*, 106:131302, 2011.

[52] D.Y. Akimov et al. ZEPLIN Collaboration. WIMP-nucleon cross-section results from the second science run of ZEPLIN-III. *Physics Letters B*, 709(1):14–20, 2012.

[53] M. Rammensee. ATLAS and CMS Collaboration. SUSY searches: Recent results from ATLAS and CMS. In *Proceedings of the Fourth Symposium on Prospects in the Physics of Discrete Symmetries, DISCRETE 2014*, 2015.

[54] G. Aad et al. ATLAS Collaboration. Search for dark matter in events with a z boson and missing transverse momentum in pp collisions at $\sqrt{s}$=8 TeV with the ATLAS detector. *Phys.Rev.D*, 90:012004, 2014.

[55] G. Aad et al. ATLAS Collaboration. Search for dark matter in events with a hadronically decaying w or z boson and missing transverse momentum in pp collisions at $\sqrt{s}$= 8 TeV with the ATLAS detector. *Physical Review Letters*, 112(4):041802, 2014.

[56] G. Aad et al. ATLAS Collaboration. Search for dark matter candidates and large extra dimensions in events with a jet and missing transverse momentum with the ATLAS detector. *Journal of High Energy Physics*, 2013(4):1–51, 2013.

[57] G. Aad et al. ATLAS Collaboration. Search for dark matter in events with heavy quarks and missing transverse momentum in pp collisions with the ATLAS detector. *The European Physical Journal C*, 72(2):1–22, 2015.

[58] V. Khachatryan et al. Search for dark matter, extra dimensions, and unparticles in monojet events in proton-proton collisions at $\sqrt{s} = 8$ TeV. *The European Physical Journal C*, 75(5):1–25, 2015.

[59] V. Khachatryan et al. Search for physics beyond the Standard Model in final states with a lepton and missing transverse energy in proton-proton collisions at $\sqrt{s} = 8$ TeV. *Physical Review D*, 91(9):092005, 2015.

[60] S. Chatrchyan et al. Search for dark matter and large extra dimensions in pp collisions yielding a photon and missing transverse energy. *Physical Review Letters*, 108(26):261803, 2012.

[61] G. Aad et al. ATLAS Collaboration. Search for dark matter candidates and large extra dimensions in events with a photon and missing transverse momentum in pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Physical Review Letters*, 110(1):011802, 2013.

[62] J. Goodman et al. Constraints on dark matter from colliders. *Physical Review D*, 82(11):116010, 2010.

[63] P.J. Fox et al. Missing energy signatures of dark matter at the LHC. *Physical Review D*, 85(5):056011, 2012.

[64] S. Profumo, W. Shepherd, and T.M.P. Tait. Pitfalls of dark matter crossing symmetries. *Physical Review D*, 88(5):056018, 2013.

[65] R. Barbier et al. R-parity-violating supersymmetry. *Physics Reports*, 420(1):1–195, 2005.

[66] J. Hisano et al. Explosive dark matter annihilation. *Physical Review Letters*, 71(6):063528, 2004.

[67] J.L. Feng, M. Kaplinghat, and H.-B. Yu. Sommerfeld enhancements for thermal relic dark matter. *Physical Review D*, 82(8):083525, 2010.

[68] I. Cholis and D. Hooper. Dark matter and pulsar origins of the rising cosmic ray positron fraction in light of new data from the AMS. *Physical Review D*, 88(2):023013, 2013.

[69] R. Battiston. The antimatter spectrometer (AMS-02): A particle physics detector in space. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 588(1):227–234, 2008.

[70] M. Aguilar et al. AMS Collaboration. First result from the Alpha Magnetic Spectrometer on the International Space Station: precision measurement of the positron fraction in primary cosmic rays of 0.5 - 350 GeV. *Physical Review Letters*, 110(14):141102, 2013.

[71] S.W. Barwick et al. Measurements of the cosmic-ray positron fraction from 1 to 50 GeV. *The Astrophysical Journal*, 482(2):L191, 1997.

[72] O. Adriani et al. PAMELA Collaboration. An anomalous positron abundance in cosmic rays with energies 1.5 to 100 GeV. *Nature*, 458(7238):607–609, 2009.

[73] M. Ackermann et al. Fermi Collaboration. Measurement of separate cosmic-ray electron and positron spectra with the Fermi Large Area Telescope. *Physics Review Letters*, 108(1):011103, 2012.

[74] O. Adriani et al. PAMELA Collaboration. PAMELA results on the cosmic-ray antiproton flux from 60 MeV to 180 GeV in kinetic energy. *Physical Review Letters*, 105(12):121101, 2010.

[75] P. Meade et al. Dark matter interpretations of the $e^{\pm}$ excesses after FERMI. *Nuclear Physics B*, 831(1):178–203, 2010.

[76] A. Ibarra and S. Wild. Prospects of antideuteron detection from dark matter annihilations or decays at AMS-02 and GAPS. *Journal of Cosmology and Astroparticle Physics*, 2013(02):021, 2013.

[77] W.B. Atwood et al. Fermi Collaboration. The Large Area Telescope on the Fermi gamma-ray space telescope mission. *The Astrophysical Journal*, 697(2):1071, 2009.

[78] F. Aharonian et al. HESS Collaboration. Observations of the Crab nebula with HESS. *Astronomy & Astrophysics*, 457(3):899–915, 2006.

[79] J. Aleksic et al. MAGIC Collaboration. The major upgrade of the MAGIC telescopes, Part I: The hardware improvements and the commissioning of the system. *Astroparticle Physics*, 72:61–75, 2015.

[80] T.C. Weekes et al. VERITAS Collaboration. VERITAS: the Very Energetic Radiation Imaging Telescope Array System. *Astroparticle Physics*, 17(2):221–243, 2001.

[81] F. Iocco, M. Pato, and G. Bertone. Evidence for dark matter in the inner Milky Way. *Nature Physics*, 11(3):245–248, 2015.

[82] J.F. Navarro, C.S. Frenk, and S.D.M. White. A universal density profile from hierarchical clustering. *The Astrophysical Journal*, 490(2):493, 1997.

[83] F. Nesti and P. Salucci. The dark matter halo of the Milky Way, AD 2013. *Journal of Cosmology and Astroparticle Physics*, 2013(07):016, 2013.

[84] L. Pieri et al. Implications of high-resolution simulations on indirect dark matter searches. *Physical Review D*, 83(2):023518, 2011.

[85] J. Stadel et al. Quantifying the heart of darkness with GHALO - a multibillion particle simulation of a galactic halo. *Monthly Notices of the Royal Astronomical Society*, 398(1):L21–L25, 2009.

[86] J.F. Navarro et al. The inner structure of CDM haloes - III. Universality and asymptotic slopes. *Monthly Notices of the Royal Astronomical Society*, 349(3):1039–1051, 2004.

[87] R.A. Swaters et al. The central mass distribution in dwarf and low surface brightness galaxies. *The Astrophysical Journal*, 583(2):732, 2003.

[88] K. Spekkens, R. Giovanelli, and M.P. Haynes. The cusp/core problem in galactic halos: long-slit spectra for a large dwarf galaxy sample. *The Astronomical Journal*, 129(5):2119, 2005.

[89] F. Walter et al. THINGS: The HI nearby galaxy survey. *Galaxies in the Local Volume*, pages 97–104, 2008.

[90] A. Burkert. The structure of dark matter halos in dwarf galaxies. *The Astrophysical Journal Letters*, 447(1):L25, 1995.

[91] K.N. Abazajian. Production and evolution of perturbations of sterile neutrino dark matter. *Physical Review D*, 73(6):063506, 2006.

[92] M.R. Lovell et al. The haloes of bright satellite galaxies in a warm dark matter universe. *Monthly Notices of the Royal Astronomical Society*, 420(3):2318–2324, 2012.

[93] A. Burkert. The structure and evolution of weakly self-interacting cold dark matter halos. *The Astrophysical Journal Letters*, 534(2):L143, 2000.

[94] C.S. Kochanek and M. White. A quantitative study of interacting dark matter in halos. *The Astrophysical Journal*, 543(2):514, 2000.

[95] S.W. Randall et al. Constraints on the self-interaction cross section of dark matter from numerical simulations of the merging galaxy cluster 1E 0657-56. *The Astrophysical Journal*, 679(2):1173, 2008.

[96] A. Loeb and N. Weiner. Cores in dwarf galaxies from dark matter with a Yukawa potential. *Physical Review Letters*, 106(17):171302, 2011.

[97] F. Governato et al. Bulgeless dwarf galaxies and dark matter cores from supernova-driven outflows. *Nature*, 463(7278):203–206, 2010.

[98] R. Teyssier et al. Cusp-core transformations in dwarf galaxies: observational predictions. *Monthly Notices of the Royal Astronomical Society*, 429:3068, 2013.

[99] P.F. Hopkins et al. Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation. *Monthly Notices of the Royal Astronomical Society*, 445(1):581–603, 2014.

[100] M. Kuhlen, M. Vogelsberger, and R. Angulo. Numerical simulations of the dark universe: State of the art and the next decade. *Physics of the Dark Universe*, 1(1):50–93, 2012.

[101] J. Diemand et al. Clumps and streams in the local dark matter distribution. *Nature*, 454(7205):735–738, 2008.

[102] J. Conrad, J. Cohen-Tanugi, and L.E. Strigari. WIMP searches with gamma rays in the Fermi era: challenges, methods and results. *Journal of Experimental and Theoretical Physics*, 121(6):1104–1135, 2015.

[103] S. Profumo, K. Sigurdson, and M. Kamionkowski. What mass are the smallest protohalos? *Physical Review Letters*, 97(3):031301, 2006.

[104] L. Gao et al. Where will supersymmetric dark matter first be seen? *Monthly Notices of the Royal Astronomical Society*, 419(2):1721–1726, 2012.

[105] M.A. Sánchez-Conde and F. Prada. The flattening of the concentration-mass relation towards low halo masses and its implications for the annihilation signal boost. *Monthly Notices of the Royal Astronomical Society*, 442(3):2271–2277, 2014.

[106] V. Springel et al. Prospects for detecting supersymmetric dark matter in the Galactic halo. *Nature*, 456(7218):73–76, 2008.

[107] F. Acero et al. Fermi Collaboration. Fermi Large Area Telescope third source catalog. *The Astrophysical Journal Supplement Series*, 218(2):23, 2015.

[108] A. Berlin and D. Hooper. Stringent constraints on the dark matter annihilation cross section from subhalo searches with the Fermi gamma-ray space telescope. *Physical Review D*, 89(1):016014, 2014.

[109] F. Aharonian et al. HESS Collaboration. The H.E.S.S. survey of the inner galaxy in very-high-energy gamma-rays. *The Astrophysical Journal*, 636(2):777, 2006.

[110] D. Hooper and L. Goodenough. Dark matter annihilation in the Galactic Center as seen by the Fermi gamma ray space telescope. *Physics Letters B*, 697(5):412–428, 2011.

[111] A. Abramowski et al. HESS Collaboration. Search for a dark matter annihilation signal from the Galactic Center halo with HESS. *Physical Review Letters*, 106(16):161301, 2011.

[112] K.N. Abazajian and M. Kaplinghat. Detection of a gamma-ray source in the Galactic Center consistent with extended emission from dark matter annihilation and concentrated astrophysical emission. *Physical Review D*, 86(8):083511, 2012.

[113] C. Gordon and O. Macias. Dark matter and pulsar model constraints from Galactic Center Fermi-LAT gamma-ray observations. *Physical Review D*, 88(8):083521, 2013.

[114] K.N. Abazajian. The consistency of Fermi-LAT observations of the Galactic Center with a millisecond pulsar population in the central stellar cluster. *Journal of Cosmology and Astroparticle Physics*, 2011(03):010, 2011.

[115] Q. Yuan and B. Zhang. Millisecond pulsar interpretation of the Galactic Center gamma-ray excess. *Journal of High Energy Astrophysics*, 3:1–8, 2014.

[116] J. Petrovic, P.D. Serpico, and G. Zaharijas. Galactic Center gamma-ray "excess" from an active past of the Galactic Centre? *Journal of Cosmology and Astroparticle Physics*, 2014(10):052, 2014.

[117] C. Weniger. A tentative gamma-ray line from dark matter annihilation at the Fermi Large Area Telescope. *Journal of Cosmology and Astroparticle Physics*, 2012(08):007, 2012.

[118] M. Ackermann et al. Fermi Collaboration. Updated search for spectral lines from Galactic dark matter interactions with pass 8 data from the Fermi Large Area Telescope. *Physical Review D*, 91(12):122002, 2015.

[119] A. Abramowski et al. HESS Collaboration. Search for photon-linelike signatures from dark matter annihilations with H.E.S.S. *Physical Review Letters*, 110(4):041301, 2013.

[120] M.G. Aartsen et al. IceCube Collaboration. Search for dark matter annihilation in the Galactic Center with IceCube-79. *European Physical Journal C*, 75(10):1–12, 2015.

[121] S. Adrián-Martínez et al. ANTARES Collaboration. Search of dark matter annihilation in the Galactic Centre using the ANTARES neutrino telescope. *Journal of Cosmology and Astroparticle Physics*, 2015(10):068, 2015.

[122] M.G. Aartsen et al. IceCube Collaboration. Multipole analysis of IceCube data to search for dark matter accumulated in the Galactic halo. *European Physical Journal C*, 75:20, 2015.

[123] L.E. Strigari. Galactic searches for dark matter. *Physics Reports*, 531(1):1–88, 2013.

[124] M. Ackermann et al. Fermi Collaboration. Searching for dark matter annihilation from Milky Way dwarf spheroidal galaxies with six years of Fermi-LAT data. *Physical Review Letters*, 115(23):231301, 2015.

[125] H. Shapley. Two stellar systems of a new kind. *Nature*, 142(3598):715–716, 1938.

[126] D.G. York et al. SDSS Collaboration. The Sloan Digital Sky Survey: technical summary. *The Astronomical Journal*, 120(3):1579, 2000.

[127] M. Ackermann et al. Fermi Collaboration. Dark matter constraints from observations of 25 Milky Way satellite galaxies with the Fermi Large Area Telescope. *Physical Review D*, 89(4):042001, 2014.

[128] S.E. Koposov et al. Beasts of the Southern wild: Discovery of nine ultra faint satellites in the vicinity of the Magellanic Clouds. *The Astrophysical Journal*, 805(2):130, 2015.

[129] J. Aleksic et al. MAGIC Collaboration. Optimized dark matter searches in deep observations of Segue 1 with MAGIC. *Journal of Cosmology and Astroparticle Physics*, 2014(02):008, 2014.

[130] E. Aliu et al. VERITAS Collaboration. VERITAS deep observations of the dwarf spheroidal galaxy Segue 1. *Physical Review D*, 85(6):062001, 2012.

[131] E. Aliu et al. VERITAS Collaboration. Erratum: VERITAS deep observations of the dwarf spheroidal galaxy Segue 1. *Physical Review D*, 91(12):129903, 2015.

[132] A. Abramowski et al. HESS Collaboration. Search for dark matter annihilation signatures in HESS observations of dwarf spheroidal galaxies. *Physical Review D*, 90(11):112012, 2014.

[133] M.G. Aartsen et al. IceCube Collaboration. IceCube search for dark matter annihilation in nearby galaxies and galaxy clusters. *Physical Review D*, 88(12):122001, 2013.

[134] M.R. Buckley et al. Search for gamma-ray emission from dark matter annihilation in the Large Magellanic Cloud with the Fermi Large Area Telescope. *Physical Review D*, 91(10):102001, 2015.

[135] A.A. Abdou et al. Fermi Collaboration. Fermi Large Area Telescope observations of Local Group galaxies: detection of M31 and search for M33. *Astronomy & Astrophysics*, 523:L2, 2010.

[136] G.M. Voit. Tracing cosmic evolution with clusters of galaxies. *Reviews of Modern Physics*, 77(1):207, 2005.

[137] T.E. Jeltema, J. Kehayias, and S. Profumo. Gamma rays from clusters and groups of galaxies: Cosmic rays versus dark matter. *Physical Review D*, 80(2):023005, 2009.

[138] M.G. Aartsen et al. IceCube Collaboration. Searches for extended and point-like neutrino sources with four years of IceCube data. *The Astrophysical Journal*, 796(2):109, 2014.

[139] M. Ackermann et al. Fermi Collaboration. Constraints on dark matter annihilation in clusters of galaxies with the Fermi Large Area Telescope. *Journal of Cosmology and Astroparticle Physics*, 2010(05):025, 2010.

[140] M. Ackermann et al. Fermi Collaboration. Search for cosmic-ray-induced gamma-ray emission in galaxy clusters. *The Astrophysical Journal*, 787(1):18, 2014.

[141] A. Abramowski et al. HESS Collaboration. Search for dark matter annihilation signals from the Fornax galaxy cluster with HESS. *The Astrophysical Journal*, 750(2):123, 2012.

[142] A. Abramowski et al. HESS Collaboration. Erratum: Search for dark matter annihilation signals from the Fornax galaxy cluster with HESS. *The Astrophysical Journal*, 783(1):63, 2014.

[143] T. Arlen et al. VERITAS Collaboration. Constraints on cosmic rays, magnetic fields, and dark matter from gamma-ray observations of the Coma cluster of galaxies with VERITAS and Fermi. *The Astrophysical Journal*, 757(2):123, 2012.

[144] G. Jungman, M. Kamionkowski, and K. Griest. Supersymmetric dark matter. *Physics Reports*, 267(5):195–373, 1996.

[145] M.G. Aartsen et al. IceCube Collaboration. Search for dark matter annihilations in the Sun with the 79-string IceCube detector. *Physical Review Letters*, 110(13):131302, 2013.

[146] K. Choi et al. Super-Kamiokande Collaboration. Search for neutrinos from annihilation of captured low-mass dark matter particles in the Sun by Super-Kamiokande. *Physical Review Letters*, 114(14):141301, 2015.

[147] M.M. Boliev et al. Baksan Collaboration. Search for muon signal from dark matter annihilations in the Sun with the Baksan Underground Scintillator Telescope for 24.12 years. *Journal of Cosmology and Astroparticle Physics*, 2013(09):019, 2013.

[148] A. Achterberg et al. AMANDA Collaboration. Limits on the muon flux from neutralino annihilations at the center of the Earth with AMANDA. *Astroparticle Physics*, 26(2):129–139, 2006.

[149] S. Desai et al. Super-Kamiokande Collaboration. Search for dark matter WIMPs using upward through-going muons in Super-Kamiokande. *Physical Review D*, 70(8):083523, 2004.

[150] SIMBAD catalog, http://simbad.u-strasbg.fr/simbad/.

[151] A. Tamm et al. Stellar mass map and dark matter distribution in M31. *Astronomy & Astrophysics*, 546:A4, 2012.

[152] J. Han et al. Constraining extended gamma-ray emission from galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 427(2):1651–1665, 2012.

[153] T. Sjöstrand, S. Mrenna, and P. Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, 2008.

[154] C. Giunti and C.W. Kim. *Neutrino Physics and Astrophysics*. Oxford University Press, 2007.

[155] I.A. Belolaptikov et al. Baikal Collaboration. The Baikal underwater neutrino telescope: Design, performance, and first results. *Astroparticle Physics*, 7(3):263–282, 1997.

[156] M. Ageron et al. ANTARES Collaboration. ANTARES: the first undersea neutrino telescope. *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, 656(1):11–38, 2011.

[157] A. Achterberg et al. IceCube Collaboration. First year performance of the IceCube neutrino telescope. *Astroparticle Physics*, 26(3):155–173, 2006.

[158] A. Gazizov and M. Kowalski. ANIS: high energy neutrino generator for neutrino telescopes. *Computational Physics Communications*, 175:203–213, 2005.

[159] P.A. Cherenkov. Visible emission of clean liquids by action of γ radiation. *Doklady Akademii Nauk SSSR*, 2:451, 1934.

[160] J.D. Jackson. *Classical Electrodynamics*. John Wiley & Sons, 1999.

[161] J.A. Formaggio and G.P. Zeller. From eV to EeV: Neutrino cross-sections across energy scales. *Reviews of Modern Physics*, 84(3):1307, 2012.

[162] L. Radel and C. Wiebusch. Calculation of the Cherenkov light yield from electromagnetic cascades in ice with GEANT4. *Astroparticle Physics*, 44:102–113, 2013.

[163] M. Kowalski. *Search for neutrino-induced cascades with the AMANDA-II detector*. PhD thesis, Humboldt Universität zu Berlin, 2004.

[164] N. Whitehorn. *A Search for High-Energy Neutrino Emission from Gamma-Ray Bursts*. PhD thesis, University of Wisconsin-Madison, 2012.

[165] J.G. Learned and K. Mannheim. High-energy neutrino astrophysics. *Annual Review of Nuclear and Particle Science*, 50(1):679–749, 2000.

[166] R.M. Sternheimer and R.F. Peierls. General expression for the density effect for the ionization loss of charged particles. *Physics Review B*, 3:3681–3692, 1971.

[167] E.L. Visser. *Neutrinos from the Milky Way*. PhD thesis, Universiteit Leiden, 2015.

[168] B. Louis et al. The evidence for oscillations. *Los Alamos Science*, page 16, 1997.

[169] J.R. Hoerandel. On the knee in the energy spectrum of cosmic rays. *Astroparticle Physics*, 19(2):193–220, 2003.

[170] T.K. Gaisser. Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio. *Astroparticle Physics*, 35(12):801–806, 2012.

[171] M. Honda et al. Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data. *Physics Review D*, 75(4):043006, 2007.

[172] G.D. Barr et al. A three-dimensional calculation of atmospheric neutrinos. *Physical Review D*, 70:023006, 2004.

[173] G. Battistoni et al. The FLUKA atmospheric neutrino flux calculation. *Astroparticle Physics*, 19(2):269–290, 2003.

[174] R. Enberg, M.H. Reno, and I. Sarcevic. Prompt neutrino fluxes from atmospheric charm. *Physical Review D*, 78(4):043005, 2008.

[175] M.G. Aartsen et al. IceCube Collaboration. Measurement of the atmospheric ν_e flux in IceCube. *Physical Review Letters*, 110:151105, 2013.

[176] R. Abbasi et al. IceCube Collaboration. The design and performance of IceCube DeepCore. *Astroparticle Physics*, 35(10):615–624, 2012.

[177] J.P. Yanez Garza. *Measurement of neutrino oscillations in atmospheric neutrinos with the IceCube DeepCore detector*. PhD thesis, Humboldt-Universität zu Berlin, 2014.

[178] R. Abbasi et al. IceCube Collaboration. Calibration and characterization of the IceCube Photomultiplier Tube. *Nuclear Instruments and Methods A*, 618(1):139–152, 2010.

[179] R. Abbasi et al. IceCube Collaboration. The IceCube data acquisition system: Signal capture, digitization, and timestamping. *Nuclear Instruments and Methods A*, 601:294–316, 2009.

[180] M.G. Aartsen et al. IceCube Collaboration. Energy reconstruction methods in the IceCube neutrino telescope. *Journal of Instrumentation*, 9(03):03009, 2014.

[181] M. Ackermann et al. IceCube Collaboration. Optical properties of deep glacial ice at the South Pole. *Journal of Geophysical Research*, 111(D13), 2006.

[182] M.G. Aartsen et al. IceCube Collaboration. Measurement of South Pole ice transparency with the IceCube LED calibration system. *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, 711:73–89, 2013.

[183] D. Chirkin. Evidence of optical anisotropy of the South Pole ice. In *Proceedings of the 33rd International Cosmic Ray Conference*, arXiv:1309.7010, 2013.

[184] D. Heck et al. CORSIKA: a Monte Carlo code to simulate extensive air showers. 1998.

[185] H.L. Lai et al. Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions. *European Physical Journal C*, 12:375–392, 2000.

[186] C. Andreopoulos et al. The GENIE neutrino Monte Carlo generator. *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, 614(1):87–104, 2010.

[187] A.M. Dziewonski and D.L. Anderson. Preliminary Reference Earth Model. *Phys. Earth Planet. Inter.*, 25:297–356, 1981.

[188] S. Agostinelli et al. GEANT4 - a simulation toolkit. *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, 506(3):250–303, 2003.

[189] D. Chirkin and W. Rhode. Propagating leptons through matter with Muon Monte Carlo (MMC). hep-ph/0407075.

[190] D. Chirkin. Photon tracking with GPUs in IceCube. *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, 725:141–143, 2013.

[191] P.B. Coates. The origins of afterpulses in photomultipliers. *Journal of Physics D: Applied Physics*, 6(10):1159, 1973.

[192] O. Schulz. *The design study of IceCube DeepCore: Characterization and veto studies.* PhD thesis, Ruperto-Carola-University of Heidelberg, 2010.

[193] M.G. Aartsen et al. IceCube Collaboration. Improvement in fast particle track reconstruction with robust statistics. *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, 736:143–149, 2014.

[194] J. Ahrens et al. AMANDA Collaboration. Muon track reconstruction and data selection techniques in AMANDA. *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, 524(1):169–194, 2004.

[195] F. James and M. Roos. Minuit: A system for function minimization and analysis of the parameter errors and correlations. *Computer Physics Communications*, 10(6):343–367, 1975.

[196] N.J.A.M. Van Eijndhoven, O. Fadiran, and G. Japaridze. Implementation of a Gauss-convoluted Pandel PDF for track reconstruction in neutrino telescopes. *Astroparticle Physics*, 28(4):456–462, 2007.

[197] D. Pandel. Bestimmung von Wasser- und Detektorparametern und Rekonstruktion von Myonen bis 100 TeV mit dem Baikal-Neutrinoteleskop NT-72. Master's thesis, Humboldt-Universität zu Berlin, 1996.

[198] K. Schatto. *Stacked searches for high-energy neutrinos from blazars with IceCube.* PhD thesis, Johannes Gutenberg Universität Mainz, 2014.

[199] N. Whitehorn, J. Van Santen, and S. Lafebre. Penalized splines for smooth representation of high-dimensional Monte Carlo datasets. *Computer Physics Communications*, 184(9):2214–2220, 2013.

[200] B. Efron. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, pages 1–26, 1979.

[201] S. Euler. *Observation of oscillations of atmospheric neutrinos with the IceCube Neutrino Observatory.* PhD thesis, RWTH Aachen, 2014.

[202] M.G. Aartsen et al. IceCube Collaboration. Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data. *Physical Review D*, 91(7):072004, 2015.

[203] T. Neunhöffer. Estimating the angular resolution of tracks in neutrino telescopes based on a likelihood analysis. *Astroparticle Physics*, 25(3):220–225, 2006.

[204] R. Abbasi et al. IceCube Collaboration. An improved method for measuring muon energy using the truncated mean of dE/dx. *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, 703:190–198, 2013.

[205] P. Speckmayer et al. The toolkit for multivariate data analysis, TMVA 4. *Journal of Physics: Conference Series*, 219(3):032057, 2010.

[206] J.R. Quinlan. Induction of decision trees. *Machine learning*, 1(1):81–106, 1986.

[207] Y. Freund and R.E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.

[208] A.N. Kolmogorov. Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, 4:83–91, 1933.

[209] N. Smirnov. Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, pages 279–281, 1948.

[210] M. Larson. Simulation and identification of non-Poissonian noise triggers in the IceCube neutrino detector. Master's thesis, University of Alabama, 2013.

[211] J. Braun et al. Methods for point source analysis in high energy neutrino telescopes. *Astroparticle Physics*, 29(4):299–305, 2008.

[212] G. Punzi. Comments on likelihood fits with variable resolution. 2004, arXiv:physics/0401045.

[213] M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, 27(3):832–837, 1956.

[214] E. Parzen. On estimation of a Probability Density Function and mode. *The Annals of Mathematical Statistics*, 33(3):1065–1076, 1962.

[215] B.W. Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.

[216] G.R. Terrell and D.W. Scott. Variable kernel density estimation. *The Annals of Statistics*, pages 1236–1265, 1992.

[217] G.J. Feldman and R.D. Cousins. Unified approach to the classical statistical analysis of small signals. *Physical Review D*, 57(7):3873, 1998.

[218] A. Tamm, E. Tempel, and P. Tenjes. Visible and dark matter in M31 - II. A dynamical model and dark matter density distribution. 2007, arXiv:0707.4374.

[219] M. Ackermann et al. Fermi Collaboration. Constraining dark matter models from a combined analysis of Milky Way satellites with the Fermi Large Area Telescope. *Physical Review Letters*, 107(24):241302, 2011.

[220] J. Han et al. Constraining extended gamma-ray emission from galaxy clusters. 2012, arXiv:1201.1003.

[221] M.G. Aartsen et al. IceCube-Gen2: a vision for the future of neutrino astronomy in Antarctica. 2014, 1412.5106.

[222] M.G. Aartsen et al. Letter of Intent: The Precision IceCube Next Generation Upgrade (PINGU). 2014, 1401.2046.

Acknowledgements

The acknowledgements have always been my favourite part of a PhD thesis to read (so perhaps it is not surprising that I have left academia after doing my PhD), and now I finally get the opportunity to write my own. There are so many people who contributed in one way or another to either the actual work done during my PhD (which is reported here in this thesis), the many things I learned during the past six years or the fun that I had while doing it all, that I am sure that I will forget some of them, but I will do my best.

First and foremost I would like to thank both of my supervisors, Elisa Bernardini and Alexander Kappes, for taking me on and guiding me through this process. Even though for both of you dark matter was not directly related to your own work, you took an interest and we have definitely had some great discussions during the past years. Alexander, thank you for your help during the first year when I was still trying to figure out the life of a PhD student, and for getting me involved in the direction reconstruction research (even though it unfortunately did not lead to any great improvements, it was a very interesting topic nonetheless). Elisa, though you've let me work very independently (which to be honest fits much better with my personality than being micro-managed), you always backed me when it was necessary and also always reminded me of the importance of my analysis, which was definitely needed every now and then. Thanks a lot for that!

Before starting a PhD, it is hard to get a real feeling of what kind of group you'll be in and what kind of collaboration you'll be in, but in both cases I would say that I was lucky. The IceCube group at DESY had quite a number of people, all working on different things, which means that you could usually find an expert in just about anything, and there were always many and varied questions and pieces of advice during talks and rehearsals. The group is too large to mention everyone by name (I am sure to forget some!), but special thanks to Christian for being such a colourful and enthusiastic (retired) group leader.

The IceCube collaboration as a whole is small enough to know (almost) everyone but large enough to (again) have experts on a lot of different topics, which was great. I am happy that I got the opportunity to join a number of collaboration meetings (especially the meeting at CERN, where everyone hung out at the CERN restaurant terrace until late at night, and the meeting in Madison, where I shared a house with fellow PhD students and got to give a plenary presentation after not sleeping all night due to an allergic reaction..). It is much easier to drop someone a quick email with questions, or ask them for a favour, when you've had a couple of beers with them before, and the meetings were always great opportunities to talk to my fellow dark matter searchers (Carlos and Carsten, our great group leaders, James who was so enthusiastic about his new-born son that he had a picture of him in his first collaboration meeting talk, Jan L who answered all my questions about dwarf galaxies and spent a lot of time in helping me figure out what was going wrong in my analysis, and many others). Thanks everyone for the good times and the ideas.

During my PhD I have also been IceRec maintainer for a few years, which was a great opportunity to learn software skills (which I still use daily) and get to know a different group of people in IceCube. Emanuel, thanks for being so enthusiastic about this and giving me the idea that I could do it, even though I am definitely not a programmer myself, and Alex and Don, thanks for welcoming me and answering all my many questions (and again, the beers during collaboration meetings, and bowling!).

One of the things that I valued most during my PhD were the friendships with people who were going through the same process, who would just understand all the frustrations that are involved, how exciting the final unblinding is, and everything in between. Kathrin, you are definitely the prime example of this as we went through most of it together (and definitely had a lot of times where both of us needed to vent..), thanks for the great times that we had at all the block courses, schools and now visits to Berlin or the Netherlands! All of the Zuethan boys, thanks for great memories in Madison and Chicago (still one of my favourite holidays) and whatever that tiny tiny village in Germany was called, thanks for letting me bring and cook vegetables every now and then.. Angel, you are probably one of the nicest guys I ever met, thanks for letting me practice my terrible Spanish on you (by now I can actually say more than just 'ola amigo, como esta', I promise). David, thanks for being my panic buddy, I am still missing all your terrible jokes! Emanuel, thanks for all your great stories and for showing all of us that there's more important things than the PhD. Hape, thanks for always being so completely cool and relaxed! Juan Pablo, thanks for being a sounding board for all my analysis-related trouble, for introducing me to Mexican music and for considering me as one of the guys, except for when you asked me once again whether I was pregnant yet.. (and again the terrible jokes!). And many others that I am just forgetting here.. doing a PhD is definitely something special, and I am glad that all of you were there to do it together.

Finally, some thanks to the people who were not directly involved in my PhD, but still know a lot more about dark matter and IceCube than they probably would like to.. my friends and family in the Netherlands, thanks for all your interested questions, for visiting us in the not-so-faraway Berlin (it was great to have so many of you stay over at our place!) and also for not talking about physics at all! Christian and Jule, Faruk and Farah, and Kathi and Matthias, thanks for all the great times we had in Berlin! My parents, thanks so much for your support, it was great that you could be there in Berlin for the defense and tell the story of how I already said as a two-year-old (when my mother had her own PhD defense) that I wanted to do a PhD.. and finally, thanks to Jesse, for everything you've done, all your support and questions and hugs and chocolate, and for this amazing adventure together in Berlin.

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Utrecht, den 02.06.2017 Meike de With